



UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

Dottorato di Ricerca in Biologia Applicata

XXV Ciclo

Curriculum in Fisiologia Vegetale

**LIPOXYGENASE PATHWAYS IN MARINE DIATOMS:
IDENTIFICATION, OCCURRENCE AND REGULATION OF
LIPOLYTIC ACYL HYDROLASEs (LAHs)**

Tutor

Dott. Angelo Fontana

Dottoranda

Dott.ssa Maria Grazia Adelfi

ABBREVIATIONS

<i>T. pseudonana</i>	<i>Thalassiosira pseudonana</i>
<i>P. tricornutum</i>	<i>Phaeodactylum tricornutum</i>
<i>P. multistriata</i>	<i>Pseudo-nitzschia multistriata</i>
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
<i>H. sapiens</i>	<i>Homo sapiens</i>
<i>P. delicatissima</i>	<i>Pseudo-nitzschia delicatissima</i>
<i>S. marinoi</i>	<i>Skeletonema marinoi</i>
<i>S. costatum</i>	<i>Skeletonema costatum</i>
<i>T. rotula</i>	<i>Thalassiosira rotula</i>
<i>P. arenysensis</i>	<i>Pseudo-nitzschia arenysensis</i>

ABSTRACT

Diatoms are eukaryotic microalgae broadly present in freshwater and marine ecosystems. They are responsible for about 40% of the marine primary production and 20% of global primary production.

The organic material formed through these processes is at the bases of the marine food chain and, in this way, supports the entire life on the planet. About the physiological and ecological studies on diatoms, however, only in recent years has begun to understand the molecular aspects. In analogy with higher plants and brown algae, diatoms possess lipoxygenase able to produce oxygenated compounds of poly-unsaturated fatty acids. In diatoms, these molecules, generically called *oxylipins*, show an unusual structural variability that seems to depend on genetic diversity. It is believed that the oxylipins have functions of intraspecific or interspecific chemical signal, however, these studies have always been limited by the lack of knowledge of the molecular mechanisms that regulate the biosynthesis of these compounds. Based on functional and tracer studies, it has been proposed that the process leading to the formation of oxylipins is due to a galactolipid hydrolyzing activity that, like lipolytic acyl hydrolases (LAHs) of land plants, is also capable of converting phospholipids. This doctoral study has focused the molecular identification of lipolytic acid hydrolases (LAHs), in three ecologically relevant diatoms: *Pseudo-nitzschia multistriata*, *Pseudo-nitzschia arenysensis*, and *Skeletonema marinoi*.

TABLE OF CONTENTS

ABBREVIATIONS.....	3
ABSTRACT.....	4
1. INTRODUCTION.....	8
1.1 Diatoms: General Characteristics.....	8
1.2 Cell Cycle and cell wall biogenesis	14
1.3 Ecological Importance.....	17
1.4 Model Organism: GENUS <i>Pseudo-nitzschia</i>	21
1.5 Model Organism: <i>Skeletonema marinoi</i>	24
1.6 Molecular Biology of Diatoms and development of molecular tools....	25
1.7 Primary and Secondary Metabolism: Biosynthesis and role of Primary and Secondary Metabolites in terrestrial and marine organisms.....	29
1.8 Oxylipins and Lipoxygenase pathway in plants.....	35
1.9 “Chloroplast transit peptide” for proteins that targeting the chloroplast in plants.....	38
1.10 Lipoxygenase pathway in diatoms.....	40
1.11 <i>Transit diatom peptide</i> for plastid proteins transport	47

1.12 The family of Lipolytic Acyl Hydrolases (LAHs).....	51
1.13 Aim of thesis.....	56
2. MATERIALS AND METHODS.....	58
2.1 Growth condition.....	58
2.2 RNA extraction and reverse transcription.....	58
2.3 Primer design	59
2.4 Semiquantitative RT-PCR.....	61
2.5 Electrophoresis on agarose gel.....	62
2.6 qRT- PCR condition	62
2.7 Primer specificity and efficiency.....	63
2.8 Mathematical model for relative quantification in qRT- PCR experiments and analysis of data.....	64
2.9 Genomic DNA extraction.....	65
2.10 Amplification of genomic DNA of <i>P. multistriata</i> with primer designed on transcript of LAH gene.....	65
2.11 Bioinformatic tools for sequences analysis.....	66
2.12 Oxylipins analysis.....	67
3. RESULTS.....	69

3.1 Search for lipolytic acyl hydrolase (LAH) homologs in the diatoms genomes using as queries genes encoding proteins of LAH family in plants.....	69
3.2 Sequence analysis in order to predict the presence of a putative "TRANSIT PEPTIDE" for the proteins that targeting in the chloroplast.....	81
3.3 Validation of internal controls (reference genes) for quantitative gene expression study (qPCR) of <i>LAH</i> gene (GENE OF INTEREST <i>GOI</i>)	85
3.4 Gene expression analysis of putative <i>LAH</i> gene in <i>P. multistriata</i>	94
3.5 Gene expression analysis of putative <i>LAH</i> gene in <i>P. arenysensis</i>	102
3.6 Gene expression analysis of putative <i>LAH</i> gene in <i>S. marinoi</i>	104
3.7 Characterization of LOX pathway and oxylinin profiling in <i>P.</i> <i>multistriata</i>	106
3.8 Characterization of LOX pathway and oxylinin profiling in <i>P.</i> <i>arenysensis</i>	111
3.9 Characterization of LOX pathway and oxylinin profiling in <i>S.marinoi</i>	114
4. DISCUSSION.....	117
5. CONCLUSION AND FUTURE PERSPECTIVES.....	129
6. REFERENCES.....	132

1. INTRODUCTION

1.1 Diatoms: General Characteristics

Diatoms (class *Bacillariophyceae*) are a group of unicellular algae that colonize widely the oceans. An important characteristic of these organisms is the ability to generate a silica cell wall, finely structured, known as *frustule*. The wall is composed of two parts, coarsely defined by the shape of a box (*ipoteca*) with the lid (*epiteca*), slightly overlapping. The bases of the two parties (*epivalva* and *ipovalva*) are joined by the elements called *epicingulum* and *ipocingulum*, which together form the belt (Figure 1). Although this structure is composed of various elements, they are closely joined together. The cytoplasm results fully protected, so that any exchange of material with the environment must be through appropriate pores or cracks. The frustule may have several silicate or chitinous appendices, which are species-specific and tightly regulated genetically (Round *et al.*, 1990) and used for the taxonomic classification.

Diatoms are divided into *centric* or *pennate* according to the shape of the frustule. The centric diatoms have a radial symmetry, while the pennate have an elongated shape with bilateral symmetry compared to the plane of the valves (Figure 2). Among the pennates there is a further subdivision according to the presence or absence of *raphe*, a polysaccharide structure

probably used for the movement of the cell on solid surfaces (Van Den Hoek *et al.*, 1997). In fact, the species that are provided of *raphe* are often benthic, while the planktonic species, both centric and pennate t have only a relic vestige of this structure (*araphidinae*). Like other eukaryotic organisms capable of carrying out photosynthesis, diatoms are equipped with plastids inside of which there are complexes of proteins and pigments important for the capture of light (Falciatore and Bowler, 2002). Diatoms are golden-brown color due to the presence of the accessory carotenoid pigment *fucoxanthin* that masks the green of *chlorophyll a* and *c*. *Fucoxanthin* and *chlorophyll* are held together within the antenna complex in order to capture the light from the Fcp proteins (*fucoxanthin chlorophyll a/c-binding proteins*). Fcp proteins are proteins homologous to Cab (*Chlorophyll a / b-binding protein*) of green algae and higher plants and, in analogy with the latter, their synthesis is activated by light (Leblanc *et al.*, 1999; Siaut, 2007). A distinctive feature of diatoms is the presence of four membranes around their plastids that is likely due to a different evolutionary origins.

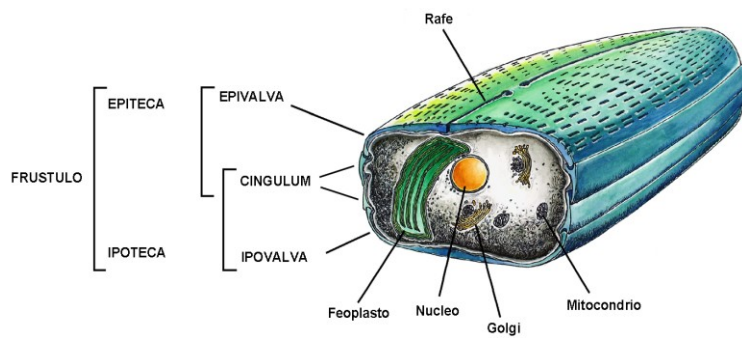


Figure 1: Schematic diagram of a *pennate* diatom (Mower and Bowler, 2002).

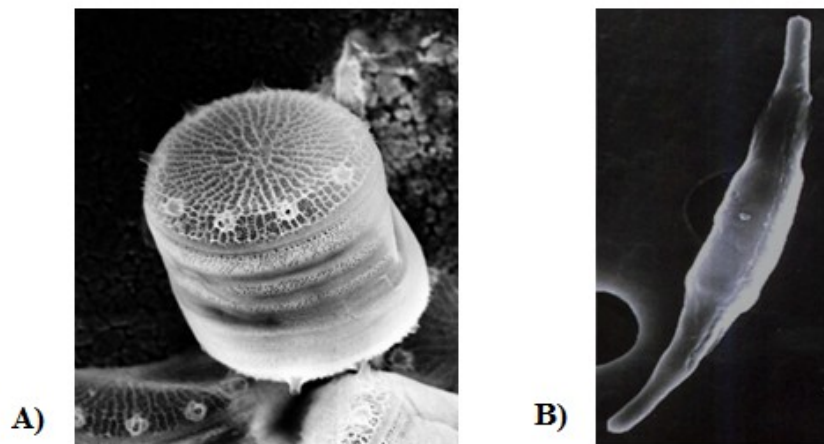


Figure 2: Pictures of the centric diatom *Thalassiosira pseudonana* (A) and the pennate diatom *Phaeodactylum tricornutum* (B).

Phylogenetic analyzes have clearly indicated that the chloroplasts of most algae and higher plants are derived from a process of primary endosymbiosis occurred at least 1.5 billion years ago from a photosynthetic bacterium (very similar to current *cyanobacteria*) and eukaryotic unicellular heterotroph. Instead diatom chloroplasts might be the result of a second event of endosymbiosis occurred about 1 billion years later in which an eukaryotic alga, probably a red alga, was incorporated into a second eukaryotic heterotrophic cell (Bhattacharya, 1995) (Figure 3).

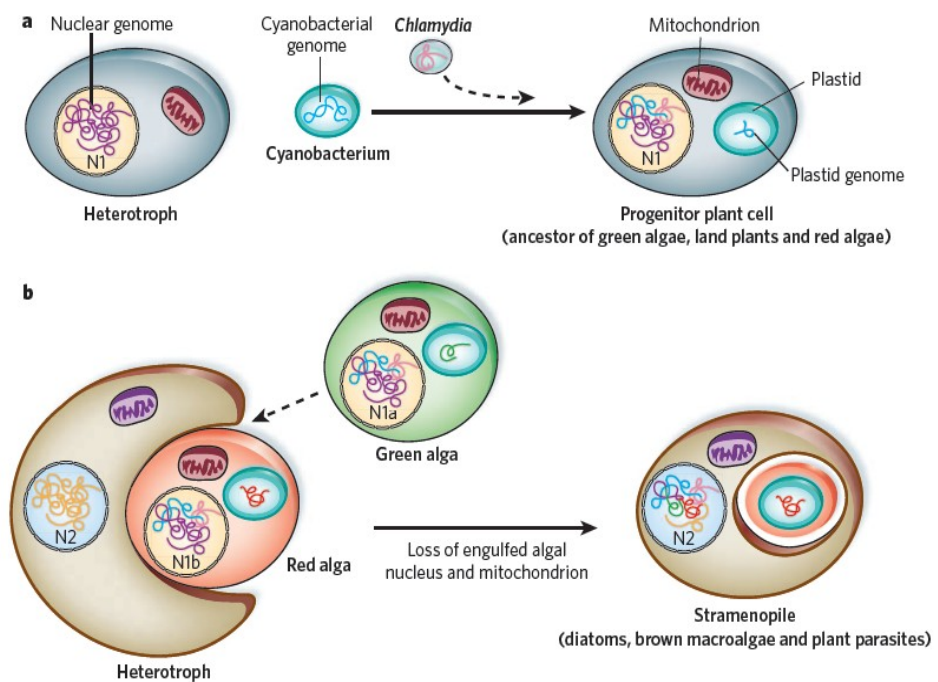


Figure 3: Representation of the origin of diatom plastids through sequential primary (a) and secondary (b) endosymbioses. (Armbrust, *Nature* 2009).

Phylogenetic studies indicate that diatoms diverge early from the lineage that gave rise to green algae and higher plants. According to this, the recent genome sequencing of two species of diatoms has shown that these organisms have unique characteristics. The most striking is definitely that half of their genes encode proteins that are related to those of animals instead of plants (Figure 4) (Armbrust *et al.*, 2004; Bowler *et al.*, 2008). Furthermore, the diatoms seem to possess unique metabolic cycles, such as the presence of a complete urea cycle, uncommon in algae and plants, which makes these organisms unique within the photoheterotrophic eukaryotes.

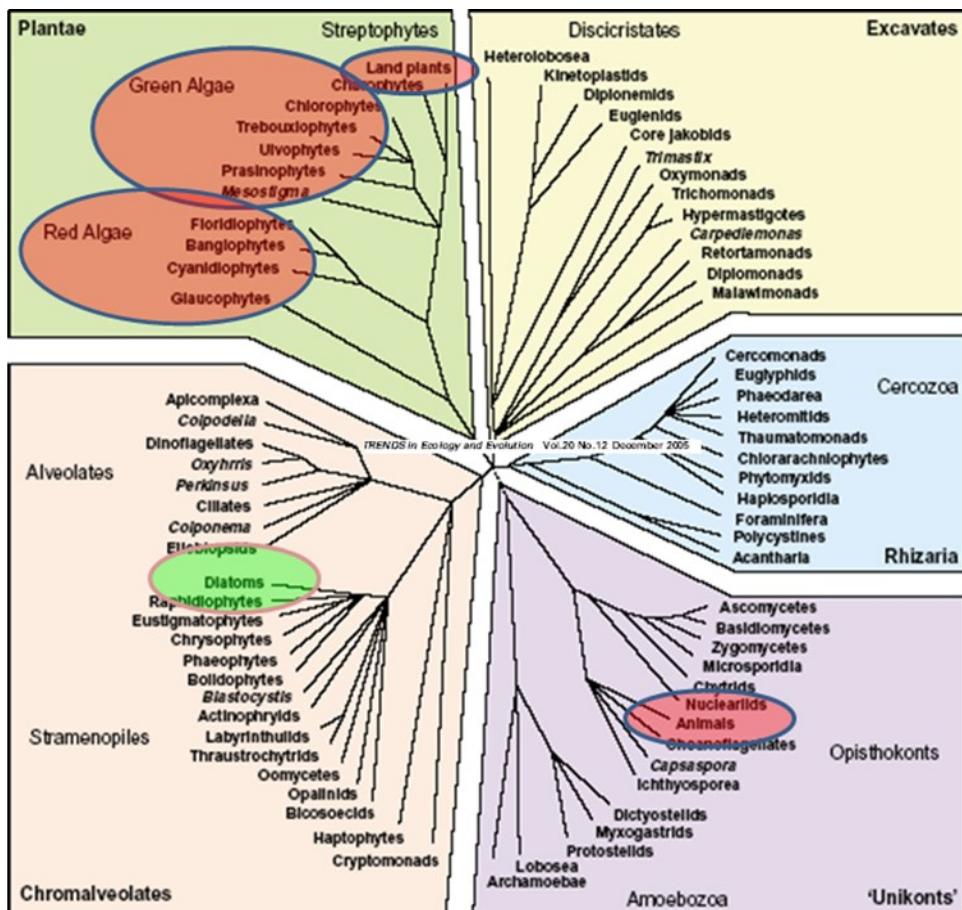


Figure 4: Tree of Eucaryotes (Keeling *et al.*, 2005, *TRENDS in Ecology and Evolution*)

1.2 Cell Cycle and cell wall biogenesis

Vegetative cells of diatoms are diploid. In a normal cycle of asexual division two daughter cells originates from a mother cell. Each valve of the parental cell becomes the *epiteca* of the daughter cell (Pickett-Heaps *et al.*, 1990) and each daughter cell will generate a new *ipoteca* (Figure 5). Before the division, the cell elongates, pushing the *epiteca* away by the *ipoteca*, and the nucleus divides through a mitosis "open". After the protoplast was divided by the invagination of the plasma membrane, each daughter cell generates a new *ipoteca*. This structure, which covers one half of the cell, is commonly generated through the production of polarized vesicles known as silica deposition vesicle (SDV). The construction of the *ipoteca* provides that the pattern of silica is wrapped by an organic matrix which prevents the dissolution. Once the entire structure is generated, the organic structure is poured out, after which the two daughter cells can be separated. The design of the wall is reproduced from one generation to another, implying the presence of a strong genetic control.

The process of formation of the frustule is poor known, however, many studies are underway to explore the molecular mechanisms behind this incredible process. In most of the diatoms, the process of wall formation causes, for each mitotic division, the reduction in the size of one of the two

daughter cells. The regeneration of the size takes place through sexual reproduction, followed by the formation of *auxospore*. Gametogenesis normally occurs when the size of the cells were reduced by 30-40%. This is considered the threshold value. The resulting male and female gametes unite to create a diploid *auxospora* that is larger than the parents. Then, the new cell goes against processes of asexual division until an appropriate signal does not induce again gametogenesis. Sexual reproduction involves several mechanisms, described in detail by Mann (1993). Centric diatoms in the sexual division is always oogama with flagellated male gametes. Among the pennate diatoms there is much more variety, and we can find cases of *anisogamia*, *isogamia* and *automixis*. In some species, sexual reproduction occurs less than once a year is limited to a short period (minutes or hours), and involves only a small number of vegetative cells in a population (Mann, 1993). When cells are exposed to unfavorable growth conditions can sometimes be induced to reproduce sexually. It has been reported that changes in the conditions of light, temperature, nutrient availability and salt concentration can induce a change in the mode of reproduction by asexual to sexual (Armbrust and Chisholm, 1990).

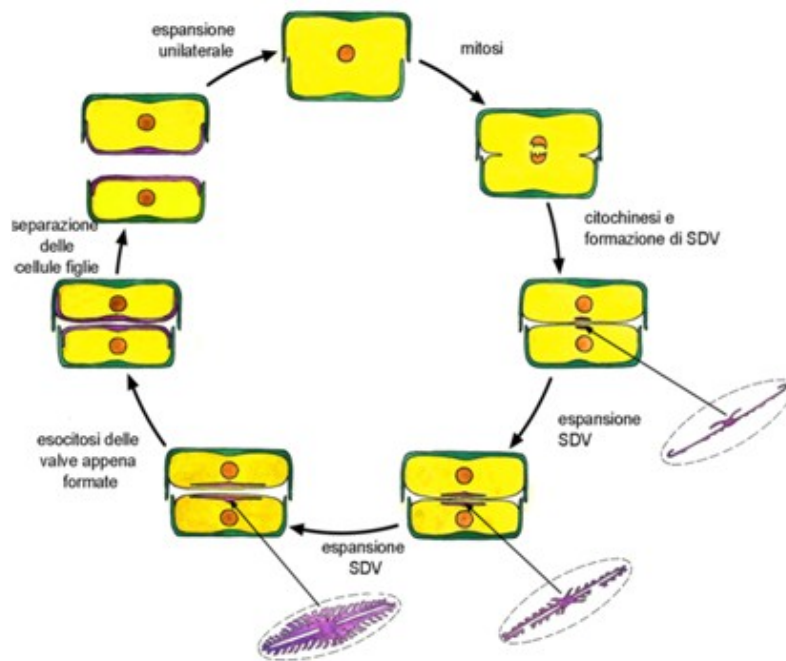


Figure 5: Scheme of mitotic cell division of a diatom.

1.3 Ecological Importance

Oceans cover about two-thirds of the Earth's surface and are thus essential components of the global ecosystem. Approximately 45-50% of primary production in the world comes from marine resources (Field *et al.*, 1998). It was calculated that diatoms account for 20-25% (Van Den Hoek *et al.*, 1997) of global primary productivity and the role of these protists is comparable to that of all forests on Earth (Smetacek, 1999) (Figure 6).

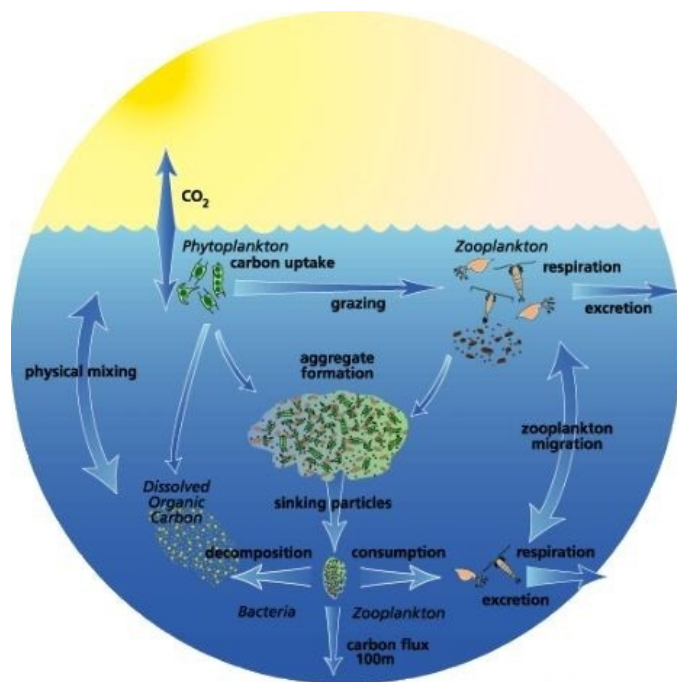


Figure 6: The carbon cycle in the ocean.

For this reason, the diatoms are always regarded as the most representative organisms of phytoplankton and predominant components at the base of the food chain of aquatic ecosystems (Figure 7).

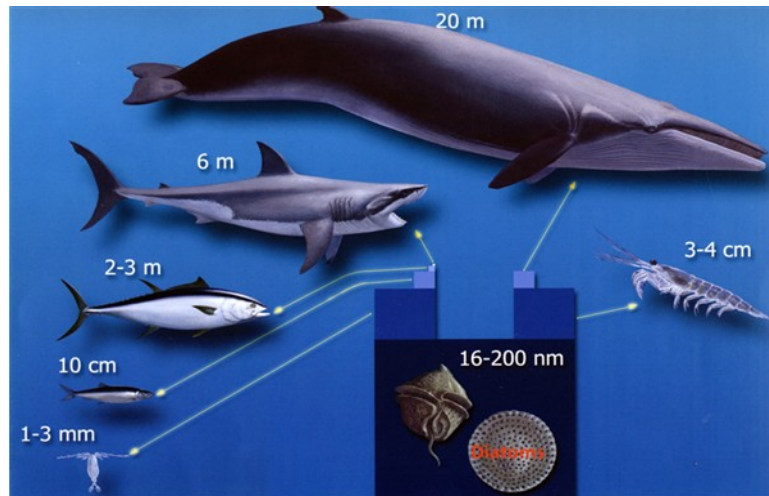


Figure 7: Marine trophic chain.

Diatoms are present virtually in any type of environment in which there are water and light. Until now 250 kinds of diatoms have been classified in 100,000 species (Norton *et al.*, 1996; Van Den Hoek *et al.*, 1997), which cover about three orders of magnitude in size and with an incredible variety of forms. The most studied species are the smallest ones (5-50 μm), which tend to be very abundant in early spring and autumn, when the concentrations of nutrients are not limiting, and when the light intensity and duration of the day are optimal for photosynthesis. When the nutrients decrease, these species tend to aggregate and precipitate below the area irradiated by the

light. Giant diatoms (that can reach 2-5 mm in size) are ubiquitous in the oceans and their distribution is less subject to seasonal variability. Their cell wall predominates in sediments of the oceans and, for this reason play a key role in biogeochemical cycles of silica (Kemp *et al.*, 2000). In addition to planktonic species there are also benthic diatoms that grow on sediments, rocks or other plants. Diatoms are the most abundant algae in the waters of the Antarctic and the Arctic, where the living conditions are extreme. On warm oligotrophic waters is possible, however, find associations between symbiotic nitrogen-fixing bacteria, cyanobacteria and diatoms (Villareal, 1989), which contribute significantly to determine the total amount of nitrogen in the ecosystem.

Many hypotheses have been advanced about the reasons for such a large and lasting ecological success, but still there is no clear answers. As shown by recent analysis (Vardi *et al.*, 2006), a possible explanation may be the presence of a sophisticated system of perception of environmental stimuli, which would be the prerequisite for the activation of the mechanisms of acclimatization. Even the peculiar cell wall has likely contributed to the success of diatoms in the course of evolution. In fact, the choice of silica as a component of the wall provides an energy advantage compared to the more "expensive" use of polysaccharides in other microalgae (Werner, 1977).

Furthermore the silica wall seems to be also an effective defense against predators (Smetacek, 1999).

Given their need for the construction of frustule, diatoms greatly influence the cycle of silicon in the sea. During diatom blooms (flowering), silica dissolved in the oceans reaches such levels that are out of range of standard analytical techniques (Round *et al.*, 1990). The following "rain" of frustules at the end of the bloom form the organic silica deposits on the ocean floor. Huge fossil deposits of diatomite (silica residues of frustules of dead diatoms) illustrate the success of this class of algae in the course of thousands of years.

1.4 Model Organism: GENUS *Pseudo-nitzschia*

Pseudo-nitzschia H. Peragallo 1899 is one of the most common kind of diatoms in the marine phytoplankton. It is present in the polar, temperate, subtropical and tropical regions worldwide (Hasle, 1964, 1965a, b, 1972a; Kaczmarska *et al.*, 1986; Fryxell *et al.*, 1991). *Pseudo-nitzschia* grow even in regions with little iron, probably due to their ability to produce ferritin, a protein that stores iron within cells (Marchetti *et al.*, 2009). The species included in the genus *Pseudo-nitzschia* are heterothallic (ie sexual reproduction occurs between clones belonging to opposite mating types) and sexual reproduction can only take place in a relatively large cell size window (Davidovich and Bates, 1998; Kaczmarska *et al.*, 2000, Amato *et al.*, 2005; Chepurnov). About 30 species of *Pseudo-nitzschia* have been described and 12 of these produce the toxin domoic acid (Moestrup, 2004). This toxin acts as an analogue of the neurotransmitter glutamate inducing apoptosis of neuronal cells. The domoic acid is responsible for the syndrome ASP (Amnesic Shellfish Poisoning) whose main symptom, found both humans and animals, is amnesia (Mos, 2001). The biological, ecological, biogeographical and biochemical interest in the species belonging to this genus has increased steadily since 1987, when an intoxication event following a bloom involved more than a hundred people in Canada. That was the first event leading to awareness that diatoms produce this toxin

(Bates *et al.*, 1989). The species of *Pseudo-nitzschia* blooms with a strong seasonality along the annual cycle in the gulf of Naples (Ribera d'Alcala *et al.*, 2004). Among *Pseudo-nitzschia* species, *Pseudo-nitzschia multistriata* and *Pseudo-nitzschia arenysensis* (Figure 8) are two species commonly found in the Mediterranean Sea for which extensive information has been collected regarding the life cycle. The heterothallic mode of reproduction makes them interesting species for the development of genetic approaches, allowing the possibility to control breeding in the laboratory.

P. multistriata was described for the first time by Takano (1993) in southern Japan as *Nitzschia multistriata* and subsequently assigned to the genus *Pseudo-nitzschia* (Takano, 1995). It has been shown that *P. multistriata* produce domoic acid (Sarno & Dahlman 2000; Orsini *et al.* 2002), the species is however only moderately toxic. According to the distribution pattern illustrated by Hasle (2002), *P. multistriata* is present near the tropics, and is generally observed in temperate waters.

P. arenysensis is a pennate, chain-forming diatom responsible for blooms in coastal and oceanic waters (Hasle, 2002). A considerable amount of genetic diversity has been recently reported among *Pseudo-nitzschia delicatissima*-like isolates and distinct cryptic species have been described within what has been identified as the *P. arenysensis* complex (Lundholm *et al.*, 2006; Amato *et al.*, 2007). Furthermore, *P. arenysensis* has been also

reported as one of the diatom species capable to produce oxygenated fatty acid derivatives (d'Ippolito *et al.*, 2009).

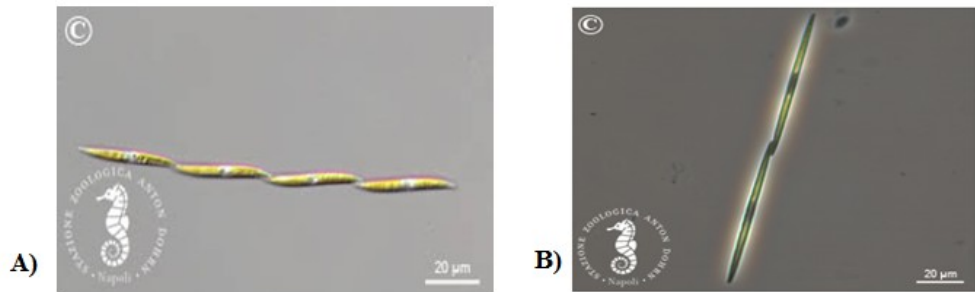


Figure 8: Images of *P. multistriata* (A) and *P. arenysensis*(B).

1.5 Model Organism: *Skeletonema marinoi*

The diatom *Skeletonema marinoi* can exist as solitary cells or form chains of more than 20 cells. The reasons why *S. marinoi* and some other species of diatoms form chain-like colonies are not clear, but other organisms (bacteria and *Phaeocystis*) can start to grow in filaments or colonies after sensing cues from grazers (Long *et al.* 2007) (Figure 9). Each cell is 2–12 μm and contains one or two large chloroplasts. The species is common in the Adriatic Sea where, in winter, is responsible for the yearly maximum of phytoplankton biomass. These blooms affect negatively copepod reproduction. The effect is ascribed to production of short chain aldehydes which interfere with the normal development of the eggs (Ianora *et al.* 2004, cited as *Skeletonema costatum*).

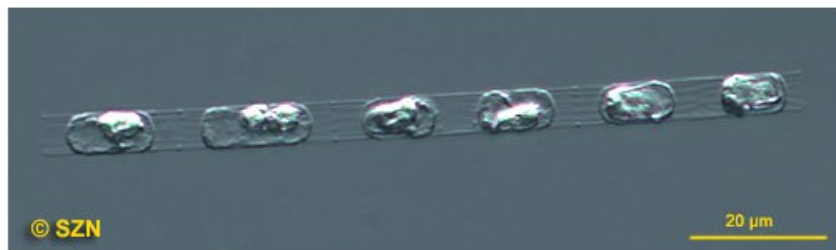


Figure 9: Image of *S.marinoi*.

1.6 Molecular Biology of Diatoms and development of molecular tools

Despite the great ecological interest about diatoms, many of the molecular issues of this group of organisms have yet to be discovered. This is essentially due to the limited availability of molecular techniques to support this research. Until a few years ago, the gene sequences of diatoms in databases were not more than 50 and any type of molecular analysis required the development of specific protocols. Nowadays diatoms are well established in terms of genomic and transgenic capabilities. There are currently four available diatom genome sequences for *Thalassiosira pseudonana* (<http://genome.jgi-psf.org/Thaps3/Thaps3.home.html>), *Phaeodactylum* *tricornutum* (<http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>), *Fragilariopsis cylindrus* (<http://genome.jgi-psf.org/Fracy1/Fracy1.home.html>) and *Pseudo-nitzschia multiseries* (<http://genome.jgi.doe.gov/Psemu1/Psemu1.home.html>). Furthermore, the transcriptomes of *P. multistriata*, *P. arenysensis* and *S. marinoi* (<http://genome.jgi.doe.gov/genome-projects>; <http://marinemicroeukaryotes.org>) have been recently completed by the Laboratory of Ecology and Evolution of Plankton (LEEP) at Zoological Station “Anton Dohrn” where I did part of my PhD work.

Sequenced genome sizes in diatoms vary from 27 Mbp to 300 Mbp, and GC content is close to 50% (Armbrust *et al*, 2004; Bowler *et al*, 2008). Most of the current molecular studies on diatoms have been performed on pennates species such as *P. tricornutum* and *Cylindrotheca. fusiformis*, although both of them do not possess significant ecological importance or specific biotechnological value.

The development of committed molecular tools is crucial to study the biology of these species. One of these is the technology of quantitative PCR (q-PCR). Real-time PCR technology has removed many of the difficulties associated with quantitative studies of gene expression (Bustin SA, 2010). This technique offers a robust tool for precisely quantifying changes in gene expression. The development of an effective protocol for the study of gene expression using the technique of qPCR can be a valuable tool for the study of the biology of diatoms from the molecular point of view. In particular, it can be particularly advantageous in those species that so far allow little manipulations, such as the diatoms of the genus *Pseudo-nitzschia*. Selection of an appropriate normalization method is crucial for reliable quantitative gene expression results (Freeman *et al*, 1999; Bustin SA, 2002). A study of the expression of genes commonly defined as “housekeeping genes” (HKGs), has been conducted in the diatom *P. tricornutum* (Siaut, 2007). The results obtained from the analysis carried out by q-PCR have shown that most of the

genes assumed to be constitutive were in fact regulated by light. The only gene that appeared to be expressed in a fairly stable manner was for the gene for histone H4 (Siaut *et al.*, 2007). Application of RT-qPCR has proven particularly fruitful in the study of marine phytoplankton, illuminating transcriptional responses to physical stressors (Rosic *et al.*, 2010a,b), nutrient limitation (Davis *et al.*, 2006; Moseley *et al.*, 2006; Berg *et al.*, 2008; Davis and Palenik, 2008; Stuart *et al.*, 2009; Whitney *et al.*, 2011; Wurch *et al.*, 2011; Bender *et al.*, 2012), and the diel cycle (Whitney *et al.*, 2011; Bender *et al.*, 2012), as well as highlighting the modulation and activity of many metabolic pathways (Moseley *et al.*, 2006; McGinn and Morel, 2008a; Mock *et al.*, 2008; Bender *et al.*, 2012). Studies on reference genes have been undertaken for *Nannochloropsis sp* (Cao *et al.*, 2012), the dinoflagellate *Prorocentrum minimum* (Ruoyu *et al.*, 2012), the antarctic ice alga *Chlamydomonas sp* (Liu *et al.*, 2012), the marine diatom *Ditylum brightwellii*. (Ruoyu *et al.*, 2012) and the brown alga *Ectocarpus siliculosus* (Le Bail *et al.*, 2008). Two reference genes, *TBP* (encoding the TATA box-binding protein) and *EFL* (encoding the translation elongation factor-like protein) were evaluated as candidates for q-RT-PCR assays in *S. costatum* and *Chaetoceros affinis* under various test conditions including growth stages, light-dark cycle phases and nutrient stresses (Kang *et al.*, 2012). The transcriptome datasets from the diatom *T. pseudonana* grown under various

conditions, were analyzed through literature-based searches for homologous reference genes, *k*-means clustering, and analysis of sequence counts (ASC) to identify putative reference genes (Alexander *et al.*, 2012).

1.7 Primary and Secondary Metabolism: Biosynthesis and role of Primary and Secondary Metabolites in terrestrial and marine organisms

The organic molecules that are produced by biochemical reactions in all living organisms is referred to as *primary metabolites*, and the intricate pathways by which organism synthesize and demolish the organic molecules, constitute the *primary metabolism*. Secondary metabolites have a restricted distributions being mostly characteristic of individual genera, species, or strain and are formed by specialized pathway usually stemming from primary metabolites. Accordingly, while primary metabolites are intimately involved in basic life processes, the secondary metabolites play as chemical mediators or regulators in specific individual, species or group of species.

Products of *secondary metabolism* tend to be synonymus of natural products, and thus are classified according to their chemical structures in terpenes, polyketides, fatty acid derivatives, alkaloids, shikimate and aminoacid derivatives (*The Biosynthesis of Secondary Metabolites*, Richard B. Herbert, 1989; *Biosynthesis of Natural Products*, P.Mannitto, 1981).

In aquatic systems there is a broader diversity of species and chemical compounds than in terrestrial ecosystems (McClintock and Baker, 2001). Plants and animals in the sea produce a variety of different, often unique, molecules that serve as protection against enemies or that are of vital

importance for signaling and reproduction. Often secondary metabolites constitute a very small fraction of the total biomass of an organism (Cannell, 1998) and the biochemical pathways that generate marine natural products are often complex and as much as specific, thus requiring the expenditure of a significant amount of metabolic energy (Cronin 2001).

Secondary metabolites and secondary pathways have evolved under the pressure of natural selection to address specific functions and therefore represent ecological or physiological effectors in the producing organism (Williams *et al.*, 1989). Marine organisms are under intense competitive pressure for space, light, and nutrients, thus it is not surprising that they have developed an array of chemical compounds to ensure survival. Furthermore, marine animals and plants live in close association with microorganisms and their body surfaces are inevitably colonized by epibiotic microbes; some marine animals harbor microorganism within their digestive tracts or even within tissues and cells. Such interactions are complex and reach from harmful diseases to symbioses of mutual benefit (Steinert *et al.*, 2000). Secondary metabolites can act as a defense strategy against unwanted colonization (infection) by microbes, or can act as a controlling factor in this host-microbe interaction. Many studies have also addressed allelopathic interactions in the marine environment and the function of secondary

metabolites as defenses against pathogens or other competing species (Ianora *et al.*, 2006).

Recently, major interest has focused on the potential role of secondary metabolites and allelopathy in controlling microalgal biology, species successions during bloom development, competition and communication within the phytoplankton. The role of secondary metabolites as information molecules in plankton environments (Steinke *et al.*, 2002) is underlined by diatom production of unsaturated aldehydes, which are involved in a stress surveillance mechanism based on fluctuations in calcium and nitric oxide levels (Vardi *et al.*, 2006). When stress conditions during a bloom and cell lysis rates increase, aldehyde concentrations could exceed a certain threshold, and possibly function as a diffusible bloom-termination signal that triggers an active cell death. Diatom-derived aldehydes may also have an allelopathic role, since they have been shown to affect growth and physiological performance of diatoms and other phytoplankton species (Casotti *et al.*, 2005).

Chemical interactions are very well known and studied in terrestrial ecosystems (Inderjit and Duke, 2003), but studies in aquatic systems have been biased by technical difficulties, mainly arising from dilution in the water medium and physical constraints such as viscosity or shear forces (Wolfe,

2000). Diatoms do possess defensive capabilities, most notably in the form of mechanical protection conferred by the silica frustule (Hamm *et al.*, 2003). It is now apparent that complex and highly evolved chemical defences may also be in operation in many species (Miralto *et al.*, 1999; Ianora *et al.*, 2006); Somewhat paradoxically, other autotrophs, primarily dinoflagellates and cyanobacteria, are renowned for producing highly toxic biomolecules (Landsberg *et al.*, 2002), some of which have been linked to reproductive failures. The ingestion of toxic dinoflagellates by female copepods has been suggested detrimental for nauplii hatching success (Nejstgaard *et al.*, 1996). Yan *et al.* (2001) demonstrated that intact cells and cellular fragments of the saxitoxin producing dinoflagellate *Alexandrium tamarense* inhibited hatching success and larval survival of the scallop *Chlamys farreri*. *Heterosigma carterae* was found to suppress egg hatching rates when fed to the copepod *Acartia tonsa* (Colin *et al.*, 2002) and severe reduction in copepod egg production rates were recorded during cyanobacterial blooms in the Baltic Sea (Sellner *et al.* 1996) Turner *et al.* (1998) proposed that ingested phycotoxins may adversely affect successive generations of herbivores. Domoic acid is the most recognised and established diatom biotoxin and has been investigated in relation to anti-herbivory functionality (Shaw *et al.*, 1997; Maneiro *et al.*, 2005). When stress conditions during a bloom and cell lysis rates increase, aldehyde concentrations could exceed a certain threshold,

and possibly function as a diffusible bloom-termination signal that triggers an active cell death. Diatom aldehydes do not induce the same acute toxicity syndromes as the more commonly recognised algal biotoxins, and this has prompted their impacts on grazers to be referred to as ‘insidious’ (Miralto *et al.*, 1999) They are however broadly cytotoxic with potential molecular targets associated with the cytoskeleton, calcium signalling and cell death pathways. These are unusual biological activities for microalgal-derived compounds which have also prompted some investigators to question whether or not the diatom aldehydes are indeed toxins in the conventional sense. Toxins or not, it is precisely their bioactivity that has prompted current interest in exploring the scope and potential of these molecules for biotechnological and pharmacological development.

Diatom aldehydes are products of lipoxygenase pathways and therefore belong to the class of molecules named oxylipins, fatty acid derivatives of polyunsaturated fatty acids. Oxylipins are important signal transduction molecules widely distributed in animals and plants where they regulate a variety of events associated with physiological and pathological processes. The family embraces several different metabolites that share a common origin from the oxygenase-catalyzed oxidation of polyunsaturated fatty acids. The biological role of these compounds has been especially studied in mammals and higher plants, although a varied and very high

concentration of these products has also been reported from marine macroalgae (Fontana *et al.*, 2007).

Marine diatoms produce a number of oxylipins, that include hydroxy acids, epoxyalcohols, -oxo acids and the above mentioned polyunsaturated aldehydes, mainly derived from eicosapentaenoic acid (EPA) and chloroplastic fatty acids (Pohnert, 2002; d'Ippolito *et al.*, 2003; Cutignano *et al.*, 2006; Fontana *et al.*, 2007a). A critical reading of the chemical structures of the molecules characterized suggests that oxylipins are the products of a series of species-specific metabolic pathways in both centric (d'Ippolito *et al.*, 2002, 2006; Cutignano *et al.*, 2006; Barofsky & Pohnert, 2009; Fontana *et al.*, 2007b) and pennate (Pohnert, 2000) diatoms, despite the marked differences in the genomic structure of the two lineages (Bowler *et al.*, 2008).

1.8 Oxylipins and Lipoxygenase pathway in plants

In plants, *phyto-oxylipins* are metabolites produced by the oxidative transformation of unsaturated fatty acids via a series of diverging metabolic pathways. These oxygenated derivatives actively participate in plant defense mechanisms, i.e. jasmonic acid (one branch of C18 polyunsaturated fatty acid metabolism) that is essential in plant defense-signaling pathways (Bl  e, 2002). The biosynthesis of oxylipins occurs in response to environmental stress (herbivory, wounding, pathogen attack, touch response, osmotic shock, drought, UV light) and developmental cues (anther dehiscence, pollen development, tuberization, storage). The key enzymes of this pathway are lipoxygenases proteins (LOXs), non-heme iron-containing dioxygenases that catalyses the oxygenation of polyunsaturated fatty acids (PUFA) to form fatty acid hydroperoxides.

Lipoxygenases carry out oxygenation at different points along the carbon chain, referred to as “positional” or “regio” specificity. The following transformations of the hydroperoxides generate a number of different metabolites that often play opposite effects. Linoleic and linolenic acid are the major polyunsaturated fatty acids in plant tissues and for this reason the most common substrates for plant LOXs (Porta & Rocha-Sosa,2002) to

produce two major products, the 9- and 13-hydroperoxy fatty acids (Siedow, 1991) (Figure 10).

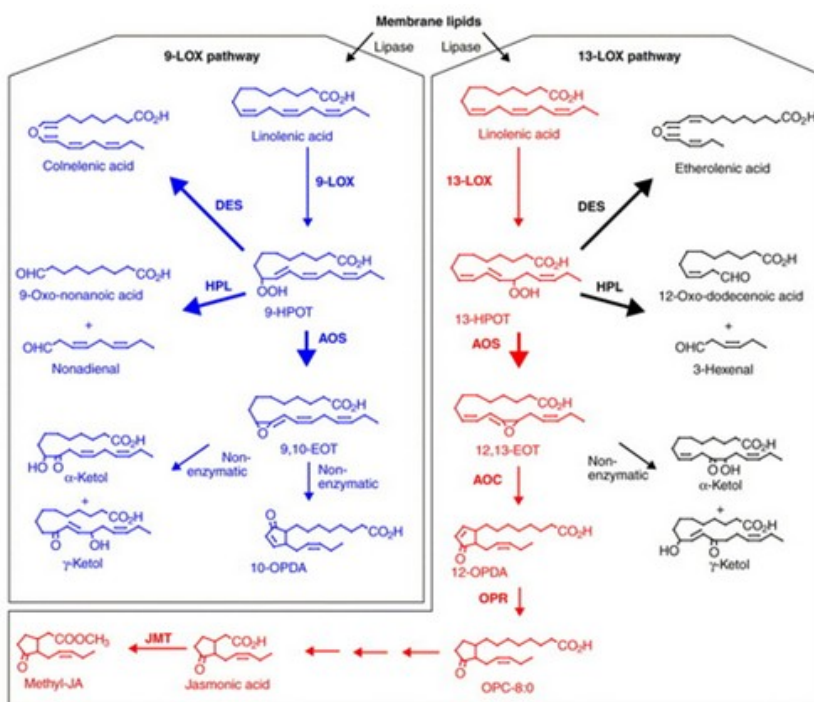


Figure 10: The 9-LOX and 13-LOX pathways for the metabolism of linolenic acid. (Howe A.G., Schilmiller L.A., 2002. *Current Opinion in Plant Biology*, Volume 5, Issue 3, pages 230-236).

In vitro, most LOXs prefer free fatty acids, though it has been shown that esterified fatty acids are also substrates for LOX in vivo (Feussner *et al.*, 2001; Stelmach *et al.*, 2001), suggesting that membrane lipids could be substrates for oxylipin biosynthesis. Phyto-oxylipin cascades also involves plastids, where unsaturated fatty acids are liberated from phospholipids or

galactolipids (present in the envelope or the thylakoids) by lipolytic acyl hydrolases (LAH).

In plants, cytosolic LOXs possess several physiological functions including progression of developmental stages (Siedow, 1991; Kolomiets *et al.*, 2001) and mobilization of storage lipids during germination (Feussner *et al.*, 2001). On the contrary, LOX isoforms found in chloroplast are responsible for synthesis of jasmonic acid (JA), which triggers gene activation during wound response. In some cases, such as the production of 9-13- hydroperoxides involved in the synthesis of C6 volatiles, require cytosolic or chloroplastic LOXs that can be in the proximity of the membrane (Porta & Rocha Sosa, 2002).

1.9 “Chloroplast transit peptide” for proteins that targeting the chroloplast in plants

Chloroplasts originated from an endosymbiotic event, in which an ancestral photosynthetic cyanobacterium was taken up by a heterotrophic host cell that already contained mitochondria (Cavalier-Smith, 2000; Rujan, T. *et al.*, 2002). This endosymbiotic process led to a massive transfer of genetic information from the endosymbiont to the emerging host nucleus. A prerequisite for the successful completion of this process was the development and establishment of protein-import machinery for chloroplast-localized polypeptides that are synthesized in the cytosol. As chloroplasts are the organelle acquired by the eukaryotic cell, several post-translational protein-targeting systems probably already existed in the host cell, for example, those belonging to mitochondria (Pfanner & Geissler, 2001), peroxisomes (Kunau, 2001) and the plasma membrane (Neuhaus *et al.*, 1998). Therefore, the arising chloroplast protein-import system had to develop unique features to ensure organelle specificity, as the correct sorting of proteins in a eukaryotic cell is essential for its functionality.

The vast majority of chloroplast proteins are synthesized as precursor proteins (preproteins) in the cytosol and are post-translationally targeted by various targeting signals to several cellular compartments (mitochondria,

peroxisomes, plasma membrane and chloroplasts). Most proteins that are committed to thylakoid membrane, the stroma and the inner envelope are synthesized with an amino-terminal extension called pre-sequence or *transit sequence*, which is proteolytically removed after import. The transit sequence is both necessary and sufficient for organelle recognition and translocation initiation. Preproteins that contain a cleavable transit peptide are recognized in a GTP-regulated manner (Svesnikova et al., 2000) by receptors of the outer-envelope translocon, which is called the TOC complex (Young *et al.*, 1999). The preproteins cross the outer envelope through an aqueous pore and are then transferred to the translocon in the inner envelope, which is called the TIC complex. The TOC and TIC translocons function together during the translocation process. Completion of import requires energy, which probably comes from the ATP-dependent functioning of molecular chaperones in the stroma (Jackson-Constan *et al.*, 2001). Then, processing peptidases in the stroma cleave the transit sequence to produce the mature form of the protein, which can fold into its native form.

1.10 Lipoxygenase pathway in diatoms

The diatoms are able to survive and proliferate in a wide range of environmental conditions by adapting their metabolism through mechanisms largely unknown. During algal blooms at sea, the hatching failure of zooplankton, due to ingestion of diatoms has been related to the presence of a family of polyunsaturated short chain aldehydes (Miralto *et al.*, 1999; d'Ippolito *et al.*, 2002; Ianora *et al.*, 2004; Pohnert *et al.*, 2002). As discussed above, different polyunsaturated short chain aldehydes have been described from different species of marine microalgae. Analysis of the antiproliferative components of extracts from the diatom *Thalassiosira rotula* resulted in the characterization of the reactive aldehydes (2E,4Z)-deca-2,4-dienal and (2E,4Z,7Z)-deca-2,4,7-trienal that inhibit egg cleavage of copepods (Miralto *et al.*, 1999). It has been already demonstrated that intact cells of *T. rotula* did not contain any of the inhibitory aldehydes and that aldehyde production is only activated after mechanical stress or cell disruption (Pohnert, 2000). Production of diatom aldehydes rely on transformation of C20 fatty acids. The process in *T. rotula* is suggested to be initiated by phospholipases that act immediately after cell damage (Pohnert, 2002). This putative phospholipase(s) might be responsible for the preferential release of free mono- and polyunsaturated fatty acids (Figure 11).

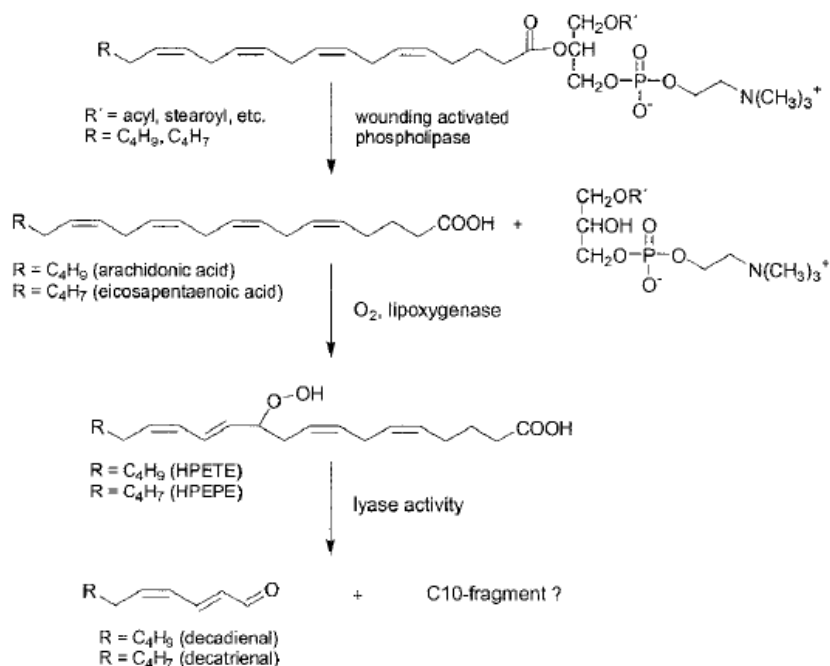


Figure 11: The activation of oxylipin-based chemical defense in the diatom *T. rotula*

(Pohnert, 2002, *Plant Physiology*, Vol. 129, pp. 103–111).

Also in the diatom *S. costatum* oxylipins are synthesized through oxidation of C20-polyunsaturated fatty acids (PUFA) (Wendel *et al.*, 1996) in response to wounding. However, d'Ippolito *et al.* (2003) showed that cells of *S. costatum* are also able to convert exogenous hexadecatrienoic acid (C16:3 N-4, HDTA) into octadienal (C8:2 N-4). This work also suggested for the first time that the biochemical pathways leading to volatile aldehydes, in analogy with land plants (Matsui *et al.*, 2000), operates through release of EPA and C16 PUFAs from glycolipids (GL), which feed the downstream

lipoxygenase (LOX)/hydroperoxide lyase (HPL) pathway. Based on functional studies, it has been proposed that the process is due to a galactolipid hydrolyzing activity that, like lipolytic acyl hydrolases (LAHs) of plants, is also capable of converting phospholipids (Figure 12). Fontana *et al.* (2007) showed that the synthesis of decatrienal occurs for LOX attack at C-11 of EPA in *T. rotula*, whereas heptadienal of *S. costatum* should derive by a similar attack at C-14. Analogously, synthesis of octadienal and octatrienal requires LOX peroxidation at C-9 of HTrA and HTA.

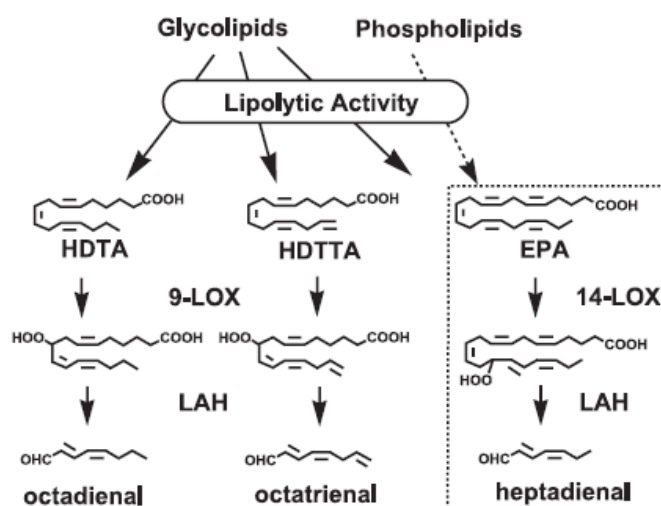


Figure 12: Hydrolysis of chloroplast-derived glycolipids leads to free forms of HDTA, HDTTA and EPA (G. d'Ippolito *et al.* (2004), *Biochimica et Biophysica Acta* 1686 100–107).

Within the ecophysiological role in marine environments, oxylipins seem to have more than one function. In addition to antimitotic activities and the consequent block of embryogenesis and induction of teratogenic effects in copepods, it was demonstrated that under physiological conditions, their synthesis varies during the phases of algal growth in a manner dependent on the density of population, suggesting therefore that these compounds may also play a role of chemical signal at the end of bloom (d'Ippolito *et al.*, 2009).

P. delicatissima produce three compounds due to 15S-LOX metabolism, namely 15S-hydroxy-(5Z,8Z,11Z,13E,17Z)-eicosapentaenoic acid (15S-HEPE), 15-oxo-5Z,9E,11E,13E-pentadecatetraenoic acid and 13,14-*threo*- 13R-hydroxy-14S,15S-*trans*-epoxyeicosa-5Z,8Z,11Z,17Z-tetraenoic acid (13,14-HEpETE) (Figure 13). Production of these compounds increases along the growth curve, with changes that precede the crash of the culture. One of the compounds, the 15-oxoacid, is formed only in the stationary phase before the collapse of the culture. The synthesis and regulation of oxylipins seem to correspond to a signaling mechanism that governs adaptation of diatoms along the growth curve until bloom termination (d'Ippolito *et al.*, 2009).

Synthesis of oxylipins depends on different enzymes mostly belonging to the family of the cytochrome P450 (CYP) superfamily. Generally the HPL converts 13-HPOT (13-hydroperoxyoctadeca-9,11,15-trienoic acid) to a hemiacetal intermediate, which is then broken into short-chain aldehydes such as green-leaf volatiles. But, for example, in *P. delicatissima*, the absence of detectable volatile aldehydes indicates that the synthesis of 15-oxoacid cannot be related to the mechanism established for plant CYP74 HPLs, which would generate downstream aldehydic products (d'Ippolito *et al.*, 2009; Hughes *et al.*, 2008; Lee *et al.*, 2008). It was been shown that the production of 15S-HEPE and 13,14-HEpETE is linearly dependent on cell density, steadily increasing with the expansion of the diatom culture, whereas the production of 15-oxoacid is constrained within the stationary phase and does not reflect a generic reaction to cell distress or death, since the compound is not produced during the declining phase when the fraction of suffering or dying cells is higher, suggesting that the activation of process may occur in response to environmental and physiological stimuli (d'Ippolito *et al.*, 2009).

From the results, it's clear that, the oxylipin composition changes among strains within the same genetically defined species, and that the LOX pathways are species-specific. These differences in LOX activities, show by species closely related, such as *P. multistriata* and *P. arenysensis*, seems to

confirm that genetic variations accumulate rapidly in the enzymes of this pathway in diatoms, or at least in the genus *Pseudo-nitzschia*.

It's was interesting also that, species related, like *Pseudo-nitzschia* species, can produce different isoforms of oxylipins, supports the functional specificity of these molecules. As well known, this functional specificity is related particularly to the role that these molecules have both as chemical signal of bloom control and defence compounds against main grazers, underlying the metabolic “plasticity” of these compounds. The modulation of LOX activities observed in *P. delicatissima*, could be due to physiological differences in the regulation of these enzymes or to the fact that positional specificity of lipoxygenases is strictly dependent on the fatty acids orientation within the enzymatic pocket (Sloane *et al.*, 1991 and 1995; Schwarz *et al.*, 2001).

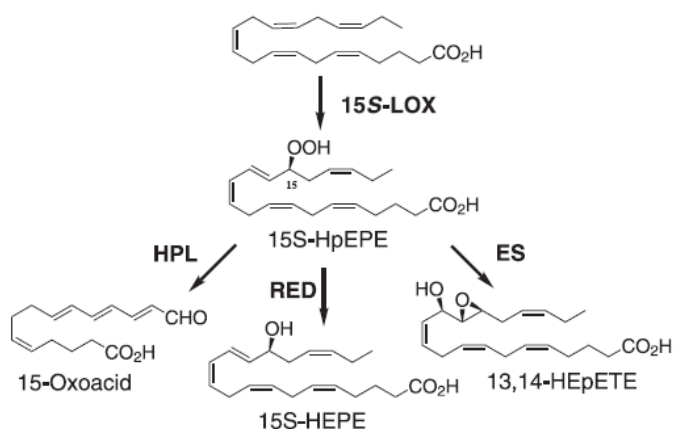


Figure 13: The 15S-lipoxygenase (LOX) pathway in *P. delicatissima* (d’Ippolito *et al.*, 2009 *New Phytologist* 183: 1064–1071).

Also in diatom *S. marinoi*, the release of oxylipins occurs transiently before the culture changes to the decline phase, providing a time-limited potential signal, supporting the hypothesis that PUA can play a role as infochemicals in mediating plankton interactions (Pohnert & Vidoudez, 2008).

1.11 *Transit diatom peptide* for plastid proteins transport

As reported above, plastids of diatoms and other chromophytic algae have four surrounding membranes. In contrast to plastids of green algae, higher plants and red algae chromophytic cells that are thought to have evolved by secondary endocytobiosis, the outermost of the four envelope membranes is studded with ribosomes and appears to be continuous with the host endoplasmic reticulum (ER) system (CER: chloroplast ER, Bouck, 1965).

Diatom thylakoid membranes contain the same classes of lipids as higher plants and green algae (Goss and Wilhelm, 2009; Lepetit *et al.*, 2011), although, in contrast to plants and green algae which have mostly neutral galactolipids, diatom thylakoid galactolipids are mainly negatively charged (Vieler *et al.*, 2007; Goss *et al.*, 2009; Lepetit *et al.*, 2011). In higher plants the most abundant galactolipids are monogalactosyldiacylglycerols (MGDG), while in diatoms the main lipids are the anionic sulphoquinovosyldiacylglycerols (SGDGs) (Goss *et al.*, 2009). Another difference is that diatom thylakoid membranes contain the phospholipid phosphatidylcholine (PC), while in plants this lipid is not present in thylakoids (Lepetit *et al.*, 2011).

All nucleus-encoded proteins of plastidic stroma in secondary algae appear to be preceded by an N-terminal extension, which consists of a *signal peptide* and a *transit peptide* (Kroth, 2002). Both domains are able to targeting plastid proteins and their capacity have been verified by in vitro experiments (Bhaya and Grossman, 1991; Lang *et al.*, 1998; Nassoury *et al.*, 2003). Preproteins are co-translationally transported into the chloroplast ER (CER) lumen via the sec protein translocators (PT). The signal peptide is subsequently cleaved off (Ishida *et al.*, 2000; Kilian and Kroth, 2003).

Two models have been described for the protein transport into diatom plastids: the “*vesicular shuttle model*” and the “*translocator model*”. In the “*vesicular shuttle model*” the intermediate preprotein still possessing a transit peptide (T) is taken up by vesicles, which shuttle within the periplastidic space (PS). After release of the preprotein into the inter-envelope space, the protein may get imported into the stroma via a protein translocator homologous to the Tic apparatus in land plant plastids (Gibbs, 1979). In the “*translocator model*” all three membranes are traversed with the aid of protein translocators. Once arrived in the stroma, the transit peptide is processed releasing the mature protein (M). (Cavalier-Smith, 2003; McFadden, 1999).

Many studies have shown that a diatom plastid targeting bipartite presequence fused to the green fluorescent protein (GFP) is sufficient for the successful transport of GFP into the plastids of the diatom *P. tricornutum* (Apt *et al.*, 2002; Domergue *et al.*, 2003; Kilian and Kroth, 2004). The characteristic feature of those presequences is the presence of a conserved motif – namely **ASAF** or **AFAP** – positioned at the cleavage site of the respective diatom signal peptides (Kroth, 2002). This motif is missing at the signal peptide cleavage site of diatom proteins putatively targeted to other destinations within the endomembrane system (e.g. into silica deposition vesicles or into the ER), thus it may represent a typical feature of plastid-targeted preproteins.

Recently, investigations have focused on the postulated translocon of the second outermost membrane (periplastidal membrane, PPM). These studies identified a symbiont specific ERAD (endoplasmic reticulum-associated degradation)-like machinery (SELMA) which has been implicated in plastid pre-protein import. This model has been described by Hempel *et al.* (2009; 2010) in *P. tricornutum*. The complex plastid of this diatom is surrounded by four membranes with the outermost one (first) being continuous with the endoplasmic reticulum (ER) of the host cell. In addition to the host-specific ERAD (ER-associated degradation) system, core components of the ERAD translocation machinery were identified in a

second version, which localizes to the periplastidal compartment (PPC) between the second and third outermost membrane of the complex plastid. According to this model, the symbiont-specific ERAD-like machinery (SELMA) mediates transport of nucleus-encoded plastid pre-proteins across the second outermost plastid membrane. Core components of SELMA are the membrane proteins sDer1-1- and sDer1-2 which putatively form the translocation channel. On the periplastidal side pre-proteins are ubiquitinated by PPC-specific Uba1 (E1 enzyme), Ubc4 (E2 enzyme) and the newly identified ubiquitin ligase ptE3P (E3 enzyme).

Subsequently, the protein is bound by the ubiquitin-dependent ATPase Cdc48 and the cofactor Ufd1, which mediate extraction of the substrate into the PPC. Following translocation ubiquitin moieties have to be removed from the substrate possibly by the newly identified de-ubiquitinating enzyme ptDUP. Finally, PPC-specific proteins are folded and processed, whereas stromal pre-proteins are transported across the third and fourth membrane presumably by a TOC/TIC-mediated mechanism

1.12 The family of Lipolytic Acyl Hydrolase (LAH)

The lipolytic acyl hydrolase (LAH), are widely distributed in animals and plants. In particular in plants are involved in processes known to operate in inducible plant defense (Shah, 2005; Andersson *et al.*, 2006; Raffaele *et al.*, 2009). Plant LAHs have been particularly examined in the context of the production of free oxylipins with prominent roles in defense responses (La Camera *et al.*, 2004) and whose biosynthesis depends on the release of fatty acids or derivatives from membrane lipids. For example, the major LAH activity induced in the tobacco hypersensitive response (HR) to pathogens as due to the accumulation of highly active patatins (Dhondt *et al.*, 2000, 2002). Their expression was tightly correlated with the appearance of cell death and activation of different branches of oxylipin metabolism.

The major known LAH gene families in *Arabidopsis thaliana* that include the patatins, DAD1-like lipases, secreted PLA2 (sPLA2) and GDSL lipases (Ryu, 2004; Matos et al, 2009) are reported in Table 1.

LAH Family in <i>A.thaliana</i>	Member name (AGI)	Confirmed subcellular localization	Transcriptional upregulation	In vitro substrate	Known functions
PATATIN	pPLA-II α (At2g26560)	cytosolic	Fungal and bacterial infections	MGDG>DGDG>PG	host cell death execution
	pPLA-III β (At3g54950)	Membrane-bound	Fungal and bacterial infections	?	?
	pPLAI (At1g61850)	Cytosolic/chloroplastic	?	MGDG>DGDG>PG	basal JA production , antifungal resistance
DAD 1	DAD 1 (At2g44810)	chloroplast	Expressed in anthers; wound-induced in leaves	PC>MGDG>TG	JA production in anthers; late JA production in wounded leaves
	DGL (PLA1-1 α -1At1g05800	Lipid bodies	Wound-induced in leaves	DGDG>PC>MGDG>TG	minor role in wound-induced JA biosynthesis?
	PLA1-I γ 1 (At1g06800)	chloroplast	Fungal and bacterial infections	Galactolipids~phospholipids ~triglycerids	?
	PLA-I γ 2 (At2g30550)	chloroplast	Fungal and bacterial infections	Galactolipids~phospholipids ~triglycerids	?
GDSL	GLIP1 (At5g40990)	Extracellular space	<i>A. brassicola</i> ethylene	Synthetic esters	Signaling; antifungal activity
	GLIP2 (At1g53940)	?	Salicylic acid	Synthetic esters	Antifungal, antibacterial resistance

Table 1: Members of LAH family in *A.thaliana*.

Most of LAH enzymes are serine hydrolases with a conserved G-x-S-x-G motif where the central Ser forms a catalytic triad with conserved Asp and His residues (Rydel *et al.*, 2003). Some features are specific to each LAH gene family, resulting in a set of LAH varying in their substrate preference and catalytic properties. A distinction can be made between broad spectrum deacylating enzymes such as *patatins* that hydrolyze both *sn-1* and *sn-2* positions of galactolipids and phospholipids but are not active on storage

triglycerides, and more specific enzymes such as sPLA2 that have an *sn*-2 positional preference on phospholipids (Matos *et al.*, 2009).

The "*patatins*" are non-specific acylhydrolases that were known first as abundant storage proteins in potato tubers (Hendriks *et al.*, 1991). The patatin-related family, pPLA (Scherer *et al.*, 2010) comprises ten genes in *Arabidopsis*, with encoded proteins falling into three subclasses (Holk *et al.*, 2002; La Camera *et al.*, 2005). There is evidence for involvement in defense responses for a member of each subclass. The subfamily of patatins-like PLA (*AtPAT*-PLA) belongs to the family of PLA in *Arabidopsis thaliana*. Between the main structural features of the members of the *A. thaliana* PLA family there are: the catalytic domains 'lipase 3' in the *AtPLA*₁, and a 'patatin' domain in the *AtPAT*-PLAs. The *AtPLA*₁-class I proteins, which bear a 'chloroplast transit peptide' domain at the N-terminus, are localized to the chloroplasts. The *AtPLA*₁-class III proteins are predicted to be localized to the mitochondria and the *AtsPLA*₂-β and -γ isoforms bear a secretory pathway signal peptide domain and are secreted into the extracellular space. pPLA-I, that is a constitutive enzyme in *Arabidopsis*, shares structural similarity with calcium-independent iPLA2α, a member of the so-called animal PNPLA protein family displaying a patatin domain and playing an important role in oxylipin biosynthesis in mammals (Kienesberger *et al.*, 2009). pPLA-I has a modular structure with leucine-rich and Armadillo

repeats in addition to the catalytical region. Proteins with LRRs include tyrosine kinase receptors, cell adhesion molecules, virulence factors, and extracellular matrix-binding glycoproteins (Enkhbayar *et al.*, 2004). They are involved in a variety of biological processes, such as signal transduction, cell adhesion, DNA repair, recombination, transcription, RNA processing, disease resistance, apoptosis, and immune response. AtPLAI is an acyl hydrolase with broad lipid substrate specificity, in fact hydrolyzes phospholipids at the sn-1 and sn-2 positions and also is more active toward galactolipids than phospholipids. Galactolipids, MGDG and DGDG, are present primarily in photosynthetic membranes inside chloroplasts (Yang *et al.*, 2007). AtPLAI has been suggested to be localized in chloroplasts (La Camera *et al.*, 2005). The linolenic acid is esterified to galactolipids, and some enzymes involved in JA biosynthesis, such as AOS and AOC, are also localized in chloroplasts (Blee, 2002; Howe *et al.*, 2002). Thus, the substrate preference and localization of AtPLAI in chloroplasts would suggest that AtPLAI affects plant functions via hydrolysis of plastidic lipids. pPLA-I was shown to hydrolyze in vitro both oxylipin- and non-oxylipin-containing galactolipids. Then there is *pPLA-IIa* encodes a cytoplasmic galacto/phospholipase that is responsible for most extractible galactolipase activity in infected leaves. The principal role of *pPLA-IIa* is to promote host cell death execution.

The subclass III of patatins, the pPLAIII β is less similar to potato tuber patatins and has not been studied in detail although it seem to be linked to the membrane and transcriptional upregulation induced by fungal and bacterial infections. The “*DADI-LIKE LAHs*” encodes a plastidial PLA1 that initiates the jasmonate synthesis and is defined like a novel family of intronless genes encoding putative lipolytic enzymes. Four isoforms have been so far characterized, all involved in fungal and bacterial infection of wound leaves.

Three of these isoforms displayed a PLA-specific activity, a PLA and galactolipase activities and broad activity towards phosphatidylcholine, galactolipids and triglycerides. The last class of enzymes belongs to GDSL lipases/esterases. This latter large family was named after their conserved motif present in the catalytic site which is distinct from the widespread GxSxG serine hydrolase motif (Oh *et al.*, 2005; Kwon *et al.*, 2009). These enzymes are localized in the periplasmic space and are involved in fungal and antibacterial resistance.

1.13 Aim of thesis

The marine diatoms are unicellular algae that play an important ecological role as major players in the global primary production (Falkowski *et al.*, 2006). These unicellular protists are able to survive and proliferate in a wide range of environmental conditions, adapting their metabolism through mechanisms largely unknown. In recent years, the research group of Institute of Biomolecular Chemistry (ICB-CNR, Pozzuoli, Naples) where I have been doing my PhD, has launched a study on identification and biosynthesis of oxylipins, oxygenated derivatives of fatty acids, which are involved in the reactions of defense and adaptation induced by abiotic and biotic stimuli in diatoms.

The subject of this work of doctoral thesis was the regulation of lipoxygenase pathway in marine diatoms by development of molecular tools to address synthesis of these compounds under physiological conditions. In particular the PhD work addressed the identification of genes encoding lipolytic acid hydrolases (LAHs), enzymes responsible for the release of fatty acids from phospholipids and glycolipids, in the ecologically-relevant diatoms *P. multistriata*, *P. arenysensis* and *S. marinoi*.

Quantitative real time polymerase chain reaction (qPCR) technique was used for the study of regulation of these genes during growth. These data

were also compared with oxylipin profiling by technique LC-MS (Liquid Chromatography-Mass Spectrometry). The results demonstrate for the first time that the synthesis of oxylipins depends on the activation of this class of proteins, thus marking a clear difference with respect to what has been so far reported in other algae.

2. MATERIALS AND METHODS

2.1 Growth condition

Cultures of *P. multistriata*, clone V.F.2.3.5, *P. arenysensis*, clone B668, and *S. marinoi*, strain FE7 (CCMP2501) were grown in f/2 medium (Guillard, 1975), incubated at 18° C under white light at approximately 60 $\mu\text{mol m}^{-2} \text{s}^{-1}$ in a 12 h photoperiod. Cells growth was monitored by cell counting using a Malassez counting chamber. Samples of three growth curves were collected during the principal growth stages: early exponential phase, mid exponential phase, late exponential phase, stationary phase. Cells were harvested by centrifugation for 15 min at 3800 rpm, washed with 1 ml of sterile sea water, aliquoted into 2 ml eppendorfs tubes, and pelleted for 3 min at 12,000 rpm at 4°C. Cells pellets were frozen instantly in liquid nitrogen and stored at -80°C, for subsequent analysis.

2.2 RNA extraction and reverse transcription

Total RNA was isolated from cells pellets at various concentration and during the various days of growth curves using High Pure RNA Isolation kit (Roche Applied Science) and the contamination of genomic DNA was eliminate with a DNase I digestion step, according to the manufacturer's

instructions. RNA concentration was determined using a NANODROP (ND 1000 Spectrophotometer) and qualitatively estimated by gel electrophoresis (1% w/v). For each sample, 500 ng of total RNA extracted was retro-transcribed with Quantiscript Reverse Transcriptase (QIAGEN), according to the manufacturer's instructions.

2.3 Primer design

For primer design of genes analyzed in this work (Table 2,3,4), specific criteria were used to standardize the reaction conditions: the melting temperature had to be between 58°C and 64°C, primer length between 19 and 24 bp. Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) was used to design primers. All primers were initially tested for PCR and the sizes of PCR products were verified on 2% (w/v) agarose gel.

NAME OF GENES	PRIMER NAME	SEQUENCES
RPS (ribosomal protein small subunit 30S)	rpsPm 2fw	5'-GTTGTCTCCAACAGAAAGGCC-3'
	rpsPm 2rv	5'-GATACGGTCGTACGAAATCTG-3'
Histone H4	H4Pm 2fw	5'-CGTGATAACATCCAGGGTATTAC-3'
	H4Pm 2rv	5'-GTGTAGGTGACGGAATCAGCG-3'
Tub A (tubulin α chain)	TubAPm fw	5'-GTTGCCGAAATCACCAGCAC-3'
	TubAPm rv	5'-GACGACATCTCCACGGTAC-3'
Tub B (tubulin β chain)	TubBPm fw	5'-CAGTGTCTTCATTGGTAACTC-3'
	TubBPm rv	5'-CATTCAAGTTGGATTACGCC-3'
TBP (TATA box binding protein)	TBPPm fw	5'-GCAATACGGAGTTCAATCCC-3'
	TBPPm rv	5'-GCTGAAAGCCTATGCGCTCG-3'
CdKA	CdKPm fw	5'-GTGCACACGGAACGGAATC-3'
	CdKPm rv	5'-CTTCAAACTCTGTGCAGTAC-3'
GAPDH (glyceraldehyde-3-phosphate dehydrogenase)	GAPDHH17F	5'-CGAGTCCACCGGTATCTTCACC-3'
	GAPDHH17R	5'-CCCATCACGAACATGGGGGC-3'
ACTIN	ACT07F	5'-CATTGCTGTTGACTTTGAAG-3'
	ACT07R	5'-CTGAATCGCTCATTTCCATAAC-3'
COPA (coatomer protein complex, subunit alpha)	COPA F1	5'-GATGCAGTTGTTGAACCG-3'
	COPA R1	5'-GGCAAAGCCATCTTCCAC-3'
LAH (lipid acyl hydrolase)	LAHPm 1FW	5'-TTATGCTGGATGGTGGGGGG-3'
	LAHPm 6 RV	5'-CACTCGTCTCAGATTAAATC-3'
	LAHPm 3 FW	5'-ATGATCTCTACATCCTTCTCTCG-3'
	LAHPm 3RV	5'-GTCTTGGTTTCTTCTCTGATC-3'
	LAHPm 1 RV	5'-ACGGGTTTGAAGACAGTTGG-3'
	LAHPm 2 FW	5'-GGAAGTAGTACTGGTGCCATC-3'
	LAHPm 2 RV	5'-TCAACACCTGGTGTCTGCTCC-3'

Table 2: List of primer for *P. multistriata*

NAME OF GENES	PRIMER NAME	SEQUENCES
RPS (ribosomal protein small subunit 30S)	RPSpare fw	5'-ACACTGACAACCTCGGCTTC-3'
	RPSpare rv	5'-GTCAACGACCTCGTCAACCT-3'
Histone H4	H4Pare fw	5'-GTCTGGAAGAGGAAAAGGAG-3'
	H4Pare rv	5'-ATACGCTTGACACCACCTCGAC-3'
Tub A (tubulin α chain)	TUBAPare fw	5'-GTCTTCGAGCCAACCAACAT-3'
	TUBAPare rv	5'-GCACCTGAATCCGGTAGGAC-3'
Tub B (tubulin β chain)	TUBBPare fw	5'-GAGGGTATGGACGAGATGGA-3'
	TUBBPare rv	5'-CACTGGCATTGCTTCATTTG-3'
TBP (TATA box binding protein)	TBPPare fw	5'-CGAGGATTGAATCCGTGTT-3'
	TBPPare rv	5'-AGAACTCCCACTCCCAAGT-3'
CdKA	CDKPare fw	5'-GTTTGAAGCAGAGGACGAG-3'
	CDKPare rv	5'-AAACCCGAATCACAACGTC-3'
GAPDH (glyceraldehyde-3-phosphate dehydrogenase)	GAPDHH17F	5'-CGAGTCCACCGGTATCTTCACC-3'
	GAPDHH17R	5'-CCCATCACGAACATGGGGGC-3'
ACTIN	ACTPare fw	5'-AGATTGTCACACACCTTC-3'
	ACTPare rv	5'-GAGGCGTAAAGGAAAGGAC-3'
COPA (coatomer protein complex, subunit alpha)	COPA fw	5'-GGAATTGTTGGAAGTGAAG-3'
	COPA rv	5'-GGTGAATACAGGTTGAGGC-3'
LAH (lipid acyl hydrolase)	LAHPare FW	5'-TCGGTTACAATGAATGGGTTC-3'
	LAHPare RV	5'-CACCAGGTGCTGTCCCGACTAC-3'

Table 3: List of primer for *P. arenysensis*

NAME OF GENES	PRIMER NAME	SEQUENCES
LAH (lipid acyl hydrolase)	LAHSke FW	5'-GAGGACACCACCGATCAAGT-3'
	LAHSke RV	5'-CAACGGCCTACCAAAGACAT-3'

Table 4: List of primer for *S. marinoi*

2.4 Semiquantitative RT-PCR

All cDNA used in this thesis were tested for RT-PCR. RT-PCR experiments were carried out with cDNAs generated from 500 ng of total RNA. The reactions were conducted generally in final volumes of 25 µl: cDNA (1 µl), 1 µl oligo fw (from stock 10 µM), 1 µl oligo rv (from stock 10 µM), PCR reaction buffer with MgCl₂ 10x (Roche), dNTP 10x, Taq DNA Polymerase (0,25 U/µl) (Roche). The thermal profile of amplification varies depending on the fragment to be amplified. In general, the PCR program included a first cycle of denaturation (2 min at 94 °C), 30- cycles with the following parameters: denaturation (45 sec at 94 °C), annealing (45 sec at 58 °C), and elongation (45 sec at 72 °C). At the end of the amplification cycles are always followed 7 min at 72 °C to allow the end of the reactions of elongation. The products were checked on agarose gel.

2.5 Electrophoresis on agarose gel

The amplification products were separated on agarose gel (2% w/v) containing ethidium bromide. The samples, to which was added 1/10 the volume of a loading solution (0.5% bromophenol blue, 30% glycerol in water) were made to run at 100 Volts for about 20 minutes.

2.6 qRT- PCR condition

Real time PCR amplification was performed with a diluted cDNA (1:5) in a reaction containing a final concentration of 1 μ M for each primer and Fast SYBR Green Master mix (Applied Biosystem) in a total volume of 10 μ l. Reaction were run in a ViiA™ 7 Real-Time PCR System (Applied Biosystems). The cycling condition was: 95°C for 20 sec, 40 cycles at 95°C for 1 sec and 60°C for 20 sec, 95°C for 15 sec, 60°C 1 min, and a gradient from 60°C to 95°C for 15 min. The results were analyzed using the *ViiA™ 7 Software* and exported into Microsoft Excel for further analysis. The Δ Ct (the difference in threshold cycles for target gene and reference gene) was calculated with the formula: Δ Ct (Avg. Ct_{tgene} – Avg. Ct_{rgene}).

2.7 Primer specificity and efficiency

To generate the efficiency curves for the different primer provided the following procedure was adopted:

- cDNA: 5 dilution points. 1:5, 1:10, 1:50, 1:100, 1:500.
- Primer: Each primer was diluted to a final concentration of 1.4 μM
- SYBR green mix: FAST method was used.

MIX: Each point analyzed in triplicate consists of 5 μl of Fast SYBR Green Master mix (Applied Biosystem); 4 μl of mixture of primer (to a final concentration of 0.7 μM for each primer) and 1 μl of cDNA dilution.

The cycling condition was: 95°C for 20 sec, 40 cycles at 95°C for 1 sec and 60°C for 20 sec, 95°C for 15 sec, 60°C 1 min, and a gradient from 60°C to 95°C for 15 min.

2.8 Mathematical model for relative quantification in qRT- PCR experiments and analysis of data

The analysis of the data was performed using the Relative Expression Software Tool-Multiple Condition Solver (REST-MCS): Calculation Software for the Relative Expression in real-time PCR using Pair Wise Fixed Reallocation Randomisation Test. The mathematical model used to determine the relative quantification of a target gene in comparison to a reference genes is the equation:

$$\text{Ratio} = (E_{\text{target}})^{\Delta\text{CP}_{\text{target (control-sample)}}} / (E_{\text{ref}})^{\Delta\text{CP}_{\text{ref (control-sample)}}}$$

The relative expression ratio (R) of a target gene is calculated based on E and CP deviation of a sample versus a control, and expressed in comparison to a reference genes. (E_{target}) is the real-time PCR efficiency of target gene transcript; (E_{ref}) is the real-time PCR efficiency of a reference gene transcript; $\Delta\text{CP}_{\text{target}}$ is the CP deviation of control – sample of target gene transcript; $\Delta\text{CP}_{\text{ref}}$ is the CP deviation of control-sample of reference gene transcript. The reference gene could be a stable and secure unregulated transcript (Pfaffl, 2001; Pfaffl *et al.*, 2002).

2.9 Genomic DNA extraction

Cells of *P. multistriata* were grown in f/2 medium (Guillard, 1975), incubated at 18° C under white light at approximately 60 $\mu\text{mol m}^{-2} \text{s}^{-1}$ in a 12 h photoperiod. Cells growth was monitored by cell counting using a Malassez counting chamber. The cells were collected in an exponential phase during a growth, and 200 ml of this culture were filtered with a filter Millipore (1.2 μm , type RAWP). The pellet resulting from the filtration was subsequently used for the extraction of genomic DNA. For genomic DNA extraction the DNeasy Plant Mini Kit (QIAGEN) was used according to the manufacturer's instructions.

2.10 Amplification of genomic DNA of *P. multistriata* with primer designed on transcript of LAH gene

To validate the structure of LAH gene and the composition exon-intron, a PCR reaction was carried out with primer designed on transcript of LAH gene: LAHPm 3FW/LAHPm 3RV (that amplify a fragments of 705 bp); LAHPm 1FW/LAHPm 1RV (that amplify a fragments of 753 bp); LAHPm 2FW/LAHPm 2RV (that amplify a fragments of 1382 bp). The reactions were conducted in final volumes of 25 μl : genomic DNA (0.5 μl), 1 μl oligo fw (from stock 10 μM), 1 μl oligo rv (from stock 10 μM), PCR reaction

buffer with MgCl₂ 10x (Roche), dNTP 10x, Taq DNA Polymerase (Roche). The PCR program of amplification included a first cycle of denaturation (2 min at 94 ° C), 35 cycles with the following parameters: denaturation (45 sec at 94 ° C), annealing (1 min at 55 °C), and elongation (1 min at 72 ° C). At the end of the amplification, the cycles were always followed by 5 min at 72 ° C to allow the end of the reactions of elongation. The products were checked on agarose gel (1% w/v).

2.11 Bioinformatic tools for sequences analysis

The nucleotide sequences were edited using DNASTar Software (Lasergene, Madison, WI, USA). BLAST and BLAST2seq (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) were used for the search of regions of local similarity between protein sequences; multiple alignments of sequences were performed with CLUSTALW (<http://www.ebi.ac.uk/Tools/msa/clustalw2>). The resultant alignments files were used to construct phylogenetic trees with the program Phylogeny.fr (http://www.phylogeny.fr/version2_cgi/simple_phylogeny.cgi) that aligned the sequences with MUSCLE method (Multiple Sequence Comparison by Log-Expectation) and design the phylogenetic trees with PhyML method, that estimate maximum likelihood phylogenies from alignments of amino

acid sequences. For prediction of putative transit peptide and subcellular localization the sequences of putative LAH proteins of *P.multistriata*, *P.arenysensis* and *S.marinoi* were analyzed with the Software: *MitoProt* (<http://ihg.gsf.de/ihg/mitoprot.html>); *TargetP* (<http://www.cbs.dtu.dk/services/TargetP/>); *ChloroP* (<http://www.cbs.dtu.dk/services/ChloroP/>); *Predotar* v. 1.03 (<http://genoplante-info.infobiogen.fr/predotar/predotar.html>), *PCLR* 0.9 (<http://www.andrewschein.com/cgi-bin/pclr/pclr.cgi>), and *PredSL* (<http://bioinformatics.biol.uoa.gr/PredSL/>),

2.12 Oxylipins analysis

Culture pellets collected for oxylipin analysis were extracted with a modified method of Cutignano *et al.* (2011), adapting the method to a small volume of cells. The pellet corresponding to 100 ml of culture pelleted in eppendorf tube was suspended in 100 µl of distilled water and sonicated 1 minute. After 30 minutes, 500 µl of MeOH and 1 µg of 16-hydroxy-hexadecanoic acid (16-OH, internal standard) were added. The sample was sonicated, and 250 µl of dichloromethane was added. The suspension was vortexed and then was taken under mild agitation for 10 minutes (the suspension must be monophasic). Then 250 µl of dichloromethane and 250 µl of water were added. The sample was centrifuged at 3500 rpm for 5 minutes to promote phase

separation. The organic phase was evaporated at reduced pressure. The dry extract was methylated with ethereal diazomethane for 30 min, and evaporated under nitrogen flow. Methylated extracts were dissolved in methanol to a final concentration of 1 µg/µl and directly analysed by LC-MS. The mass spectrometry (MS) method was based on a micro-Quadrupole time-of-flight (micro-QToF) instrument equipped with an electrospray ionization (ESI) source in positive ion mode and a UV photodiode array (DAD) detector (scan range 205-400 nm) for a dual monitoring of the chromatographic runs. For ESI-QToF-MS/MS experiments, argon was used as collision gas at a pressure of 22 mbar. Chromatographic analysis was carried out on a reverse phase column (Phenomenex, C-18 Kromasil 4.6 x 250 mm, 100 Å) using a linear MeOH/ H₂O gradient 75/25 to 100/0 in 30 min with a column flow of 1 ml min⁻¹. One tenth of the column flow was channelled by a post-column split to the ESI+ (Q-ToF).

3. RESULTS

3.1 Search for lipolytic acyl hydrolase (LAH) homologs in the diatoms genomes using as queries genes encoding proteins of LAH family in plants

The search of genes that have or could have a key role in the pathway of oxylipins production was carried out on the basis of homology with other organisms, using conserved protein domains. An important contribution came from the possibility to access to the genomes of four diatom species: *T. pseudonana*, *P. tricornutum*, *F. cylindrus* and *P. multiseriata*. In addition, the laboratory where the work was carried out has generated transcriptomes for *P. multistriata*, *P. arenysensis* and *S. marinoi*.

The search in the transcriptomes of the three species of diatom was performed using genes encoding proteins belonging to the family of lipolytic acyl hydrolase (LAH) in plants. The family is divided in three classes: 1) the *patatin-related family*, pPLA, that comprises ten genes in *Arabidopsis*, with encoded proteins falling into other three subclasses, 2) the “*DADI-LIKE LAHs*” that encodes a plastidial PLA1 that initiates the jasmonate synthesis, and 3) the classes of “GDSL lipases/esterases” (Grienenberger *et al.*, 2010). Among these, only 8 gene encoding LAH proteins were used as queries for the search: 1) AT1G61850 galactolipase/phospholipase, 2) AT2G26560

PLA2A, 3) AT3G54950 patatin-like protein, 4) AT5G40990 GDSL esterase lipase 1, 5) AT2G44810 putative phospholipase A1, 6) AT1G05800 galactolipase DONGLE, 7) AT1G06800 phospholipase A1-Igamma1,8) AT1G53940 GDSL-motif lipase 2. Among the different sequences used as queries for the Blast in the genome of *T. pseudonana*, *P. tricornutum*, *F. cylindrus*, *P. multiseriis* and in the trascriptome of *P. multistriata*, *P. arenysensis* and *S. marinoi*, only one shows a significant percentage of homology with those from genome: the acyl hydrolase/galactolipase from *A. thaliana* (At1g61850, AtPLAI). In the genome of *T. pseudonana*, *P. tricornutum*, *F. cylindrus* and *P. multiseriis* the best hit obtained blasting the sequence of *A. thaliana* shows proteins involved in metabolism of lipids, with the conserved motif of patatin-like PLA (GxSxG) and the catalytic dyad with Serine and Aspartic acid. These sequences show similarities also with intracellular membrane-bound Ca²⁺-independent phospholipase A2 from *Homo sapiens* (Figure 14).

```

>jgi|Phatr2|46193|estExt_fgenes1_pg.C_chr_90211
MLVVLYCSDCLSTIIIVAPSLWTRNPHGHGFPREAKRAPRESNVISKSEFVPQMQLSPFQGSRKRAYPSS
ASRPDWTQNSPKHRLGLPTLVAILLGNISIRSHAFVPHSTRSPALVQSRIRLADQITEKEQDESVMKWNNA
SGSFGSLMLQLQKKEEALLAVNETFLEQDVVDLDPGRGEHRAEEARTAPDLKPLEEQSSQEQLATMDWDTA
KELDNTIITILSNGDRARQELQTLPLSALQKSLLDGDTTTSPEQADSTVILPLSRKAHYEERIGRDLRHLA
VSIAASTETVADWQTFCANQGGLYPVLETIREGAQHVAPLYLNTKVNNTATSMDSQGINRNPIRMESFQA
ACSACRALRDLCAISPEVSAVLTDGILRANAWEGLMRDFTCLLQYASDIDGFTRLDGMPTLSRGQSR
KEASVNEAANAVQNRFTIIRKQRDRDRLRCKLYITQLLLAVIVASDDAVAAIRKTDGLAHAVLQSSSFAR
KQQRRLWLRYPVEMAKWLNRRRRRTVTQESANLRRPFLEAANIANDLEGQVQRTANQILAAIGHNQWRPK
IPGQKGLRILCLDGGGRGMTAVSAVKCMMDAAGNGNDVADSFDLIGTSTGIIIAFLIGLRRESSAQAV
ERYNQLIKQIFVKFFLSTPLMVFTTASYDESPFMNILSTILQDKIMLDSRADPAVFLVCCLTSMKSSSTPT
HVALFRNYNVAKDELDPFTINPDKAREDLGLPLSVEHPSVQSGSYKKELSLKNAPGARLAAGSRYPGSF
RVLQRYALRASTAAPTTFKPMVMGGEMYCDGGIVASNPAAVAVHEARTLFFDPIELVSVVGTAFLQK
SAPRIGWDGIIIGQIVNSATDGEQIHILEDLGESSILGPRSSVSKTRYFRNFVIGMPDEFIDVTDPK
KLTFLRLQLTKDYMYNEPEQCAKLLQISDLLGGRDRNKQLVISKPGEPKILAFTEPIALPHMTKETKFIKCS
LVCVLAADAVSCISPFAPRLPRAANSAPKDTISRCKRSLANARTHGLHDCRINWSPLRIRGGGIPGVVKAL
FLTALRNPVLVLLMAGSSLTIIYKSKIPHAQLQGFLSFCTAVYVYFLLTVQRMVQAETRAS*

>jgi|Frac1|185101|e_gw1.6.1043.1
MRKQERETGTDSSGRQSAIDDDGIAKELDDSVSIRITNSLKDITLLELPEIYRIRFTVNDNKEEKTADL
PPLSRAEHYDGRIGRDMRLAVSVSSIDDVAEWQLFCEQNSGGGLEPIECIRDGATSIREGNEEVLS
KNNRSDHYLEVNFQAASSACRALRNLCAISDLDAAVITDGLLRANTAYTNKGEYILMEDLCTILRHADDF
KLYVTQLLLAMTASDCAVDAIRSTEGLQDVLLVYSSYVRKERRRRMMRYPGELIKSTRRQPFIEAASLK
NDLNGRIQGTANVLAAGYNEWVFKIPGQKGLRILTLDGGGRGMTSVVAMKAIVDVLGGMIEADCFDL
VIGTSTGIIIAFLVGLNRETSEQAVERYDDLIGKIFTKSAFSTPMLLFTTATYDESPFMNVLEILGDNT
MLDSRANPAVVFVAVTSKMSSTPHIALFRNYNNGGEQDPFMDPDDARENGLSVQDEAKTIRINS
YPKKEKSLDPAASGSRHPSFRVLQKHALRASTAAPTTFKPMVMGGEMYCDGGIVASNPAAIHAERNIF
PDVPIELVVSIGTGGFKEQKSEPRIGWDGIIIGQIINSATDGEQIHILEDLVLDGDTTAQGASSVSNTRYM
RNFPIELGMPDEYPIDVTDPEKLQKIKITITSEYMEQPKQQRKLQDLADILKGRSGRIRNRNRV*

>jgi|Thaps3|23984|estExt_fgenes1_pg.C_chr_90194
MAVSILWGLPFSFISTGASAHNIPAFANSKYRSIHQFKVSHQLPPSLFYTTESIEQTSSELPTSPEEDD
TASITEESTIEDNKDWISATRTLGSFLRQEDANRDPATANNRAEIEFTSFPFQENSMTSYLLKLKQRE
EDNREKNGKGLLEKDRMVGPTSRNVFRIDQKIARELDSEVFTLPSIGDGLMGDLVLEPMIEHECDGSAK
PFENTLKAKDMLRLSSPEHYEDRIGRDMRLHGVSAATIDKFWQWKLFFDEGGVLPLEECIRDGARSVE
KGGADFVDGDIEGTSSLMQHEASFAAACTACRALRDLALSCKDFAAVVTDILKREPNVVGSEFVSAL
VAEGGDELMLAMVASDRAVITLIRSTGLIEEVLECSSYAPSERFKRWIRKPLGFIKERIPSLAKTQMN
TNLLRSRLQEDANKLLAAIGHNVVWPKLPQGRGLRILSLDGGGRGIAAAVTSIRHIVEAMGGVEVCDADF
MIVGTSTGIIIAFLVGLRRESAADARIRYDTLIKRIFKVSLKPIMLATTATASYDEANLMDVLQELKDD
GMLDSRANPEVPLITAVSSKMSSTPSQLCLLRNYYGGGELNDSFCIDPIKARQLGLEHDDVEESFPST
EPDQGTIVIKCAPRTIGSRYPGSFRVTQKIALRATTAAPTFFKPLLSEELYVDGGIVASNPATAVAVHE
ARSVFPGVPLELIVSVGTGVFEEIKVPRVGVWDGVVAQILDSATDAEQVHVLELDVFGEGRTAQLRGTKM
DSTAYFRFNAIVGKPSDFPIDEIDFVRLQELCNIVDRYMAEEKQQQKQLQGLNLHPPSIFQRAFRRILTH
EK*

>jgi|Psemul|326739|estExt_fgenes1_pg.C_4580003
MSSEKRSPPGSAEQSTTTIVASGETCATSRSLRLRAWQLFCQQTGGLEPLIECIREGAASIRAIREGSAP
FTATAPGPPSVTGTITYKSGFSVKEENFQAASSACRAIRDLCAISDLASVITDGLLRANAAYKANGEEHT
LLDDLDTPSPPEEESATTIVYQESSPSSKASSKWPMAIFRERREARLRWKLVTQLLAMTCTSDSAVDAI
RSTEGKDTLLVYSSYARKERRRRWLRYPGEMIKSVLNTKRRKRGRHRRNDSSDSSEEEETEQRQ
PFIEAASLKDNLKGRIGKTANQVLAAGYNEWVFKIPGQKGLRILCLDGGGRGMTSVVAMKCLVDSLGG
MEVADCFDLVIGTSTGIIIAFLVGLNRETSEQAVERYDVLIEKIFTKSAFSTPMLLFTTASYDESPFMNV
LTEILKDNITMLDSRANPAVVFVAVTSKMSSTPHIALFRNYNNGGELPDPFMINPDDARENGLLLEN
EAKMIRTDNDYPPKKGVASAPGTITPDNGSRHPSFRLLQKYALRASTAAPTTFKPMVMGGEMYCDGGIV
ASNPSAIAIHEARTIIFPDVPIELVVSIGTGGFKEQKSEPRIGWDGIIIGQIINSATDGEQIHILEDLVLD
GTTAQKSSSVSNTCYMRFPNIPILGMPDEFIDVTDPEKLQKIKITITKAYMEEPEQKQKLQALSILKGRNG
NRF*

```

Figure 14: Protein sequences of putative LAH in *P. tricornutum*, *F. cylindrus*, *T. pseudonana* and *P. multiseriata*. Red boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

The species chosen as model organisms for my PhD project were *P. multistriata*, for which there are no experimental data on the biosynthetic pathway of oxylipins, *P. arenysensis* and *S. marinoi*, for which LAH activity has been characterized chemically and biochemically (Fontana *et al.*, 2007; d'Ippolito *et al.*, 2009). In the transcriptome of *P. multistriata*, a sequence of 3467 bp was identified and was verified by sequencing PCR fragments amplified from total cDNA. This sequence shows two start codons (ATG), one at position 186-188, resulting in an ORF (open reading frame) of 3114 bp, and one at position 345-347, resulting in an ORF of 2955 bp, while the stop codon is at position 3297-3298 bp for both ORFs. For this reason it is not possible to establish with certainty what is the ORF coding for the protein. The first ORF codify for a protein of 1038 aa (Figure 15), whereas the second ORF for a protein of 985 aa (Figure 16).

The validation of the sequence obtained from the transcriptome and the information about the structure of the gene, were later confirmed from the access, not yet public, to the genome of this strain. The results show that the complete sequence of the LAH gene is 3904 bp, and is characterized by the presence of 5 exons and 4 introns (Figure 17). In order to confirm the structure of the gene, PCR was carried out on cDNA and on the genomic DNA of *P. multistriata*, using primers introns spanning (LAHPm 1FW/LAHPm 1RV; LAHPm 3FW/LAHPm 3RV; LAHPm 2FW/LAHPm

2RV), which amplify regions of the entire genomic sequence. In the cDNA template, the primer should be a continuous exon sequence whereas in the genomic DNA the primers should be divided by an intron. In fact from amplification of genomic DNA for PCR, I got fragments of expected length (Figure 18).

LAH P.multistriata VF2.3.5

```
MFVIGNNKR D TSSDRHLRR D RIVNAKRNP VTSRWEERT QHFKDVLFF ALCMISTSF SEQCAVEAFQ IDSVGAGPRQ RHKAQEQFRT ISASTTTTTS
QSCSESAAKM QLSVVPKVA EDNYSGDSIE TETMTVAGNS TSQFLMNFNA SGSIGSLLMQ MQKKEAELRQ LNQSASLLLS EDQALKLDPN DGSSTNQEPT
ALKNNLFSWI RKTNGSAKNG FADVDAPTAS EGVNGSTAE A ISSMNDMAR ELDDSVSIRI TNSVKDITLL GPPDIYQSR SEEQNQDDDT PFRIVLVEEE
TQTETNAKNE NAVIEDKLPP LSRPEHYDRR IGRDMRHLV SIASCIDSV EWLFCQQTTP GGLAPLVECI REGAESIREG SPSLASLSLE HQYLASSPYS
SGSTTKEENF QAASSACKAI RDLCAISLDL ASVITDGLLR ANAAYKEQGE PSLMDDLCTI LRYDDELSDL LVAPRRRRRR RLLSFRRRKG KQAGQETPQP
DPSENDPLSS PTKSPKWP I GIFRRRREAR LRWKLYVTQL LLAMTCASDS AVDAIRCTEG LQDVLLVHSS YARKERRRRW MRYPGELIKS MWLKKRKLRLN
NKEASPATPE QRRQPFIEAA SLKANLRGRI TGTANQVLA IGYNEWVPKI PQQKGLRILC LDGGGSRGMT SVVAMKCLVD SLGGMEVADC FDLVVGTSTG
AIIAFLVGLN RETSEQAVER YDVLIEKIFT KSAFSTPMLL FTTASYDESP FMNVLTDLK DRMTLDSRAN PAVPLVFAVT SKMSSNPTHV ALFRNYSYSG
GELPDPFMIN PDEARENLDL PLGNEAKIIR MSNYRKKEEH VQPVDETVPV DNGSRHPGSF RLLQKYALRA STAAPTVFKP VLMGGEMYCD GGIVASNPSA
IAIHEARTVF PGVPIELVVS IGTGGFKEQK SEPKIGWDGI IGQIINSATD GEQIHILED VLGDGTTAQG KSSVSDTCYM RFNFVLGMPD EFPIDVIDPD
KLQQIKTITK AYMEEPDQKQ KLQVLADILK GRSSTRC
```

Figure 15: Protein sequence (1038 aa) of putative LAH in *P. multistriata*. Red boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

LAH P.multistriata VF2.3.5

```
MISTSFLSEQ CAVEAFQIDS VGAGPRQRHK AQEQFRTISA STTTTTSQSC SESAAKMQLS VVPLKVAEDN YSGDSIETET MTVAGNSTSQ
FLMNFNASGS IGSLLMQMQK KEAELRQLNQ SASLLSLEDQ ALKLPDNDGS STNQEP TALK NNLFSWIRKT NGSAGNGFAD VDAPTASEGV
NGSTAEAISS MNDEMARELD DSVSIRITNS VKDITLLGPP DIYQSRSEEE GNQDDDTFPR TVLVEEETQT ETNAKNENAV TEDKLPLSR
PEHYDRRIGR DMRHLSVSA SCIDSVEEWQ LFCQQTTPGGL APLVECIREG AESIREGSPS LASLSLEHQY LASSPYSSGS TTKEENFQAA
SSACKAIRDL CALSLDLASV ITDGLLRANA AYKEQGEPSI MDDLCTILRY DDELSDLLVA PRRRRRRRL RLLSFRRRKGQA GQETPQDPDS
ENDPLSSPTS KPSKWP I GIFRRRREARLRW KLYVTQLLLA MTCASDSAVD AIRCTEGLQD VLLVHSSYAR KERRRRMMRY PGLIKSMWL
KKRKLRLNKE ASPATPEQRR QPFIEAASLK ANLRGRITGT ANQVLAAGY NEWVPKIPGQ KGLRILC LDGGGSRGMTSVV AMKCLVDSL G
GMEVADC FDL VGTSTG AII AFLVGLNRET SEQAVERYDV LIEKIFTKSA FSTPMLLFTT ASYDESPFMN VLTDLKDR T MLDSRANPAV
PLVFAVTSKM SSNPTHVALF RNYNYSYGEL PDPFMINPDE ARENLDLPLG NEAKIIRMSN YRKKEEHVQP VDETVPV DNG SRHPGSFRL L
QKYALRASTA APTVFKPVL M GGEMYCDGGI VASNPSAIAI HEARTVFPV PIELVVSIGT GGFKEQKSEP KIGWDGIIGQ IINSATDGEQ
IHILEDVLG DGTTAQGKSS VSDTCYMRFN PVLGMPDEFP IDVTDPKLQ QIKTITKAYM EEPDQKQKLQ VLADILKGRS STRC.
```

Figure 16: Protein sequence (985 aa) of putative LAH in *P. multistriata*. Red boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

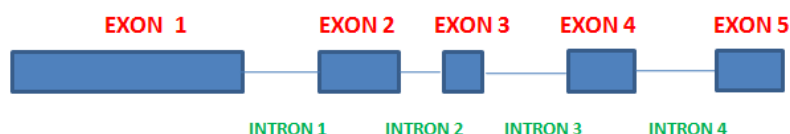


Figure 17: Structure of *LAH* gene (3904 bp) in *P. multistriata*

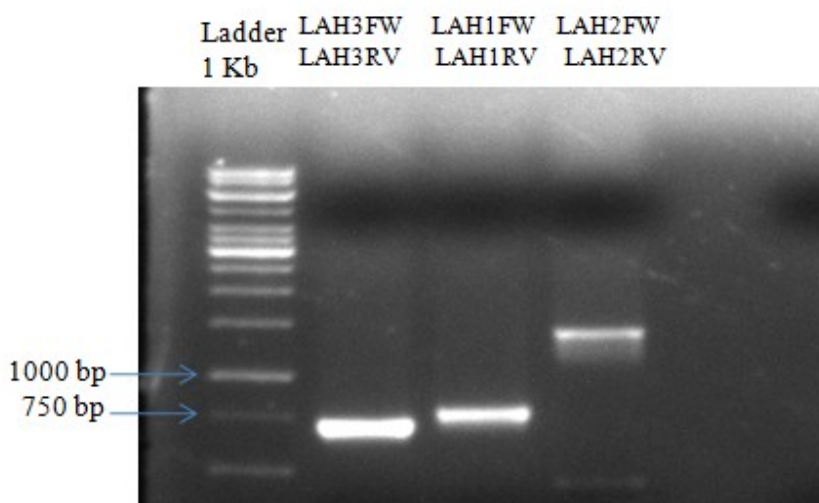


Figure 18: PCR amplification of genomic DNA of *P. multistriata* with primer: LAHPm 3FW/LAHPm 3RV (705 bp); LAHPm 1FW/LAHPm 1RV (753 bp); LAHPm 2FW/LAHPm 2RV (1382 bp). The ladder was the “Ladder 1 kb” (Fermentas).

The comparison carried out with the program ClustalW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) between the sequence identified in the transcriptome of *P. multistriata* and the sequences of *A. thaliana* clearly shows that the highest percentage of homology is with the protein acyl hydrolase/galactolipase At1g61850 (37 % homology), while the percentage of homology with other sequences is: 21 % homology with

AT2G26560 PLA2A, 15% homology with AT3G54950 patatin-like protein 6, 9% homology with AT5G40990 GDSL esterase lipase 1, 5% homology with AT2G44810 putative phospholipase A1, 4 % with AT1G05800 galactolipase DONGLE, 1 % homology with AT1G06800 phospholipase A1-Igamma1, 0% with AT1G53940 GDSL-motif lipase 2. The degree of homology between the various aligned sequences is represented in the phylogenetic tree shown in Figure 19.

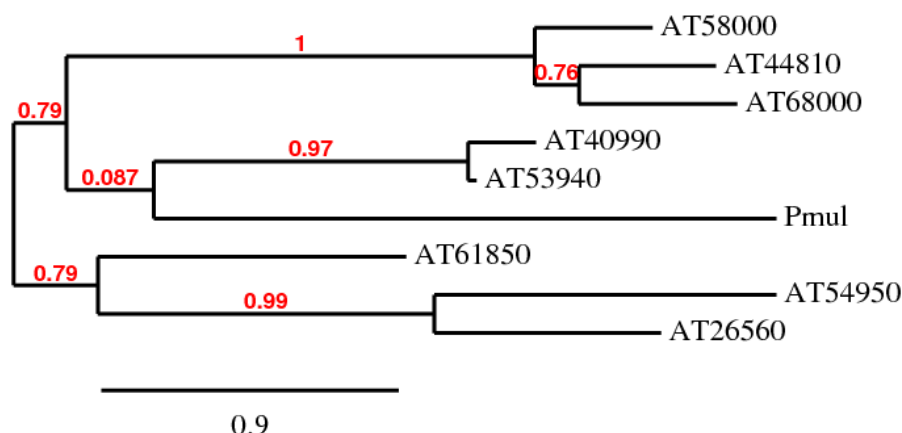


Figure 19: Phylogenetic tree based on protein sequences of *A. thaliana* (LAH family): At1g61850, At2G26560, At3G54950, At5G40990, At2G44810, At1G05800, AT1G06800, AT1G53940, and protein sequence of *P. multistriata*, obtained with *Phylogeny.fr* program (PhyML).

In the transcriptome of *P. arenysensis*, a sequence of 3143 bp, with an ORF of 2571 bp, was isolated. The corresponding protein is 857 aa (Figure 20).

LAH *P.arenysensis* B668

```

MQMQKKEAEL RLVNKSASLL FSEDQALNLD AAKSNTSASK IATTEENTAT SQKSDLFSWI RKQDGEPSGK GSTSTKAVDQ
TKIDDEIAME LDNSVSIRIT NSVKDVTILG PPEINRIKRR TGVAGESAST DNQEHEIEIG GQKNVVEEIL PTLSKSDHYD
GRIGRDMRHL AVSTASCIDS VEEWQLFCQQ STGGLEPLIE CIREGAKSVR EGGSSQTGLK YKTGLARKEE NFRMASSACK
VLRDLCALSL DLGAVITDGL LRANAAYKNN GQDTLMDDLC TILRHADDFL DIQQPASESR FRQNKREGE NEIVKPGIRS
SQSKSKWPMG FFRGRREARL RCKLYVTQLL LAMTCANDSA VDAIRSTENL QDTLLTHSSY ARKERRRRWM RYPGEKIKSM
WLTKRKIRQN LSSKTSESQK PAKKRRQPF I EAASLQNNLN GRIKGTANQV LAAIGYNEWV PKIPGQKGLR ILCIDGGGR
GMTSVIAMKC LVDSLGGMEV ADCFDMVGT STGAIIAFLV GLNLESSEKA VERYDDLIGK IFTKSTFSTP LLLFTTASYD
ESTFMNVLTLD VLGDYTM LDS RANPAVPLVF AVTSKMSSNP THIALFRNYN YNGGELPDPF LIRPDDAREN LDLPLKDEGK
SIRLNSYPKK KGIEQPAPGT ITPDNGSRHP GSFRVLKKYA LRATAAPT V FKPVLGGEM YCDGGIVASN PTAIAIHEAR
TIFPDVPIEL VVSLGTGGFT ELKSEPRIGW DGIIGQIINS ATDGEQIHHL LEDVLGDGTT AQGKPSVSNT SYMRFNPTLG
MPDEFPIDVT DPKLQQIKT ITRAYMKEPE QQRKLKIISD ILKGRSSRSR FRWFRG.

```

Figure 20: Protein sequence of putative LAH in *P. arenysensis*. Red boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

Also in this case, the comparison between the sequence identified in the transcriptome of *P. arenysensis* and the sequences of *A. thaliana* clearly shows that the highest percentage of homology is always with the protein acyl hydrolase/galactolipase Atlg61850 (50 % homology), while the percentage of homology with other sequences is: 37 % homology with AT2G26560 PLA2A, 19% homology with AT3G54950 patatin-like protein 6 , 9% homology with AT5G40990 GDSL esterase lipase 1, 2% homology with AT2G44810 putative phospholipase A1, 3 % with AT1G05800 galactolipase DONGLE, 2 % homology with AT1G06800 phospholipase A1-Igammal, 8% with AT1G53940 GDSL-motif lipase 2. The degree of homology between the various aligned sequences is represented in the phylogenetic tree shown in Figure 21.

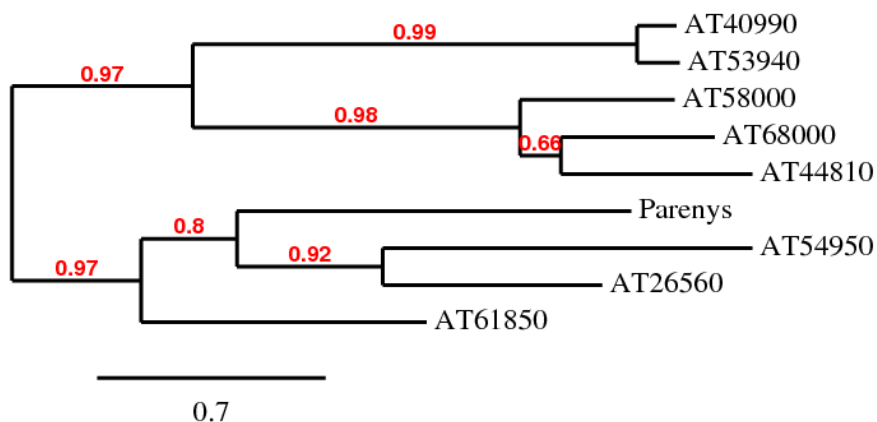


Figure 21: Phylogenetic tree based on protein sequences of *A. thaliana* (LAH family): At1g61850, At2G26560, At3G54950, At5G40990, At2G44810, At1G05800, AT1G06800, AT1G53940, and protein sequence of *P. arenysensis*, obtained with *Phylogeny.fr* program (PhyML).

In the transcriptome of *S. marinoi*, a sequence of 2748bp was isolated, with an ORF (open reading frame) that start at position 574, but it is incomplete because lacks the stop codon. (Figure 22) The comparison between the sequence identified in the transcriptome of *S. marinoi* and the sequences of *A. thaliana* shows also in this case that the highest percentage of homology is with the protein acyl hydrolase/galactolipase At1g61850 (50 % homology), while the percentage of homology with other sequences is: 39 % homology with AT2G26560 PLA2A, 16% homology with AT3G54950 patatin-like protein 6, 5% homology with AT5G40990 GDSL esterase lipase 1, 2% homology with AT2G44810 putative phospholipase A1, with AT1G05800 galactolipase DONGLE no relevant similarity, 5 % homology

with AT1G06800 phospholipase A1-Igammal, 1% with AT1G53940 GDSL-motif lipase 2. The degree of homology between the various aligned sequences is represented in the phylogenetic tree shown in Figure 23.

LAH *Skeletonema marinoi*

```

MTSHVLYAVCLLLSLPTQCTSLRSNDAFTSNYILNQPSIRKTTVLLYQIEIIDEPLNKRRISSPAKKKSKDEDTTQQVVLNDTNSKNI
QLDNNEAVEDNKOWISATRNIGSLILHLKDETKDSNYDVFGRPLNLPQIDQFGLPKDYDNNDDNGAGWQSSLASVLLKLRDEEDNRERI
FEENKGRKRNIHEDGESSKLSHNIQLDQKMKELDESYKQLSITDNLMDIQVLPNSQNSNRRNVEPKTLQGGYIGQDILRLSYPEHYKN
RIGRDMRHLAYSIAASIDQPWQWKLFFDEGGGYLPLECIRDGARSVEKGRADGDIETANSQQLLEQQQDEYSFAAACTACRALRDL
ALSKDFAAVYTDDILRVNEQWSTCYVEGEGYDCYSGGLISDILLRHANEAESFYNSRGNDRKGLRGSSRRSRTDRKGRQEARRRRCGLY
VYQLLAMVYASDKAAETLRATSGLIEAVEECSSYAPSERFRKKWYRKPLNFIGRKLGLRKTPLRKDDWLKPGKGGYQQNANKLLAGIG
HNIWTPKLPQGKGLRILCLDGGGTRGIAAYNSIGHMYKAMKGVEYCDADFMIIVGTSTGAIYGVFLVGLRRRESAADARIRYDYLKIRIFYKS
LLKPIMLATTTASYDEANLMEVLEEILKDDGMLDSRANPAYPLITAYSSKMSSTPSQLALLRNYNNGGGEMPDSFCIDPQKARKRLGLAN
DDIMEPDQSDNLPKKSTIKCAPRTGKGSRYPGSFRTQKIALRATTAAPTFFKPLLSFDELYDGGILASNPTSAYVHEARSYFPDPIE
LIYSVGTGRFEEIKYPPRVGWDGIVAQILDSATDAEQVHHAMEDVYFEGESNAKGGPQMSSTKYFRFNPLIGEPOSPFIDEIDPERLQELC
DIYDRYMDDEQKTKL

```

Figure 22: Protein sequence of putative LAH gene in *S. marinoi*. Red boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

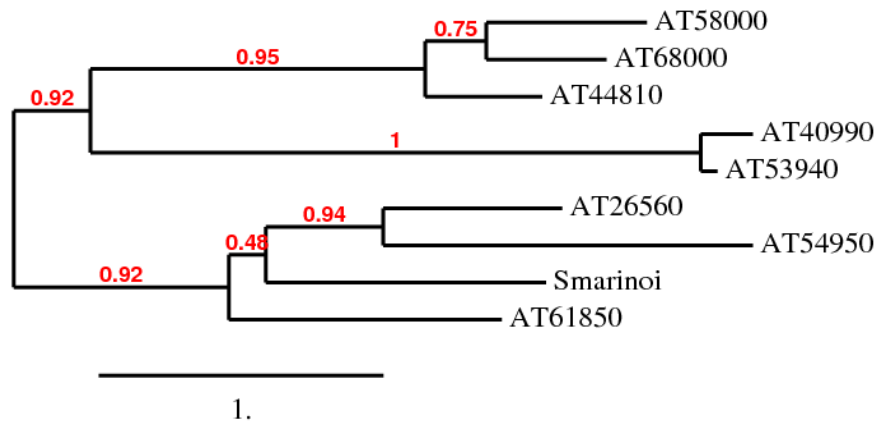


Figure 23: Phylogenetic tree based on protein sequences of *A. thaliana* (LAH family): At1g61850, At2G26560, At3G54950, At5G40990, At2G44810, At1G05800, AT1G06800, AT1G53940, and protein sequence of *S. marinoi* obtained with *Phylogeny.fr* program (PhyML).

All isolated sequences in genomes and transcriptomes of diatoms analyzed and the eight sequences used as queries for the blast and belonging

to the family of proteins LAH in plants have been used to build a phylogenetic tree. All the sequences are more closely related to the acyl hydrolase/galactolipase At1g61850, as shown previously (Figure 24). As shown in the alignment performed with ClustalW, all the sequences share the conserved hydrolasic motif GxSxG, and a catalytic dyad consist of Serine and Aspartic acid (Figure 25). These elements are critical for the LAH activity of all patatin proteins.

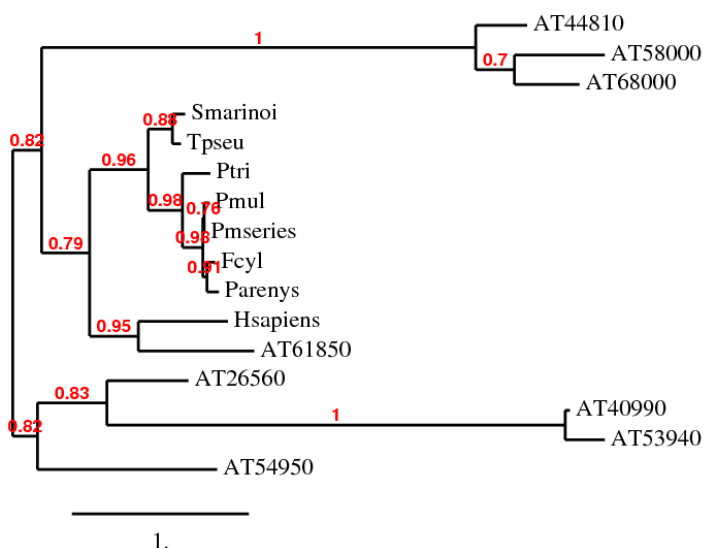


Figure 24: Phylogenetic tree based on protein sequences of *A. thaliana* (LAH family): At1g61850, At2G26560, At3G54950, At5G40990, At2G44810, At1G05800, AT1G06800, AT1G53940, sequences of putative LAH proteins in diatoms (*P. multistriata*, Pmul; *P. multiseri*, Pmseries; *F. cylindrus*, Fcyl; *P. arenysensis*, Parenys; *P. tricornutum*, Ptri; *S. marinoi*, Smarinoi; *T. pseudonana*, Tpseu) and the protein sequence of calcium-independent phospholipase A2 from *H. sapiens* (Hsapiens) obtained with *Phylogeny.fr* program (PhyML).

AT44810	IEKA--PSWMA-----TQSSWIGYVAVCQDKKE--ISRLGRRDVVISFRG-----	126
AT68000	FSKSRWSKVWS-----KNANWUMGYVAVSDDNEAT--RCRLGRRDIAIAWRG-----	218
AT58000	KNEP-----MRARWIGYVAVS--SDES--VKRLGRRDILVTFRG-----	205
Pmul	LAAIGYNEWVP-----KIPGQKGLRILCLDGGGSRGMTSVVAMKCLVDSLGLG-GMEVADCF	691
Pmseries	LAAIGYNEWVP-----KIPGQKGLRILCLDGGGSRGMTSVVAMKCLVDSLGLG-GMEVADCF	183
Fcyl	LAAIGYNEWVP-----KIPGQKGLRILCLDGGGSRGMTSVVAMKCLVDSLGLG-GMEIADCF	348
Parenys	LAAIGYNEWVP-----KIPGQKGLRILCLDGGGSRGMTSVIAMKCLVDSLGLG-GMEVADCF	657
Ptri	LAAIGHNQWRP-----KIPGQKGLRILCLDGGGSRGMTAVSAVKCMMDAAGNGNDVADSF	603
Smarinoi	LAGIGHNIWTP-----KLPGQKGLRILCLDGGGTRGIAAVNSIGHMVKAMK-GVEVCDAF	589
Tpseu	LAAIGHNVWVP-----KLPGQKGLRILCLDGGGTRGIAAVTSIRHIVEAMG-GVEVCDAF	489
AT40990	PDFIAEYANLP-----LIPPNLQPFNGNSQFAYGVNFASGGAGALVGTTFSG-----	131
AT53940	PDFIAEYANLP-----LIPAYLQPSNGKNQFPYGVSFASAGAGALVGTTFPG-----	133
Hsapiens	LALIG---YVD-----PVKGR-GIRILSIDGGGTRGVVALQTLRKIVELTQ-----	369
AT61850	LAILGENEILRRSISKGRQVPKQGLRILTMGGGTRKGLATVQILKEIEKSGS--KPIHELF	533
AT26560	VKSLTGPKYDG-----KYLHQLIHAKLGDTKLSQTLTNVVIPTFDIKHLQ-----	161
AT54950	LTRNAKGLYGG-----AGILKRVLRTGSGCCSGTAKLKKVMKESFSLETLK-----	220
AT44810	-----TATCLEWLENLR-----ATLTHLPNGPTGANL	153
AT68000	-----TVIRLEWIADLK-----DFLKPVVS-----GNGF	241
AT58000	-----TVINHEWLANLK-----SSLTPAR-----LDP	227
Pmul	DLVVGTTSTGAIIAFLVGLNRETSEQAVERYDVLIETKIFTKSAFSTPMLLFTTASYDESPF	751
Pmseries	DLVVGTTSTGAIIAFLVGLNRETSEQAVERYDVLIETKIFTKSAFSTPMLLFTTASYDESPF	243
Fcyl	DLVVGTTSTGAIIAFLVGLNRETSEQAVERYDDLIGKIFTKSAFSTPMLLFTTASYDESPF	408
Parenys	DMVVGTTSTGAIIAFLVGLNLESSEKAVERYDDLIGKIFTKSTFTSTPLLLFTTASYDESTF	717
Ptri	DLICGTSTGAIIAFLIGLRRESSAQAVERYNQLIKQIFVKPFLSTPLMVFTTASYDESPF	663
Smarinoi	DMIVGTSTGAIVGFLVGLRRESAADARIRYDVLIKRIFFVKSLLKPIMLATTTASYDEANL	649
Tpseu	DMIVGTSTGAIVAFVGLRRESAADARIRYDTLIKRIFFVKSLLKPIMLATTTASYDEANL	549
AT40990	-----LVINLRTQLN-----NFKKVEEMLRSKLGD	156
AT53940	-----MVINLKSQLN-----NFKKVEKLLRSTLGE	158
Hsapiens	-----KPVHQLFDYICGVVS-----TGAILAFMLGLF	395
AT61850	DLICGTSTGSMIAIALGVKLMLEQCEEIYKNLGKLVFAESVPKDNEAASWREKLDQLYK	593
AT26560	-----PTIFSSYEVKN-----HPLKDATLADIAIST	187
AT54950	----DTLKPVLIPCYDLKS-----SGPFLFSRADLETDGY	252

Figure 25: Alignment with the program ClustalW, between the sequences of *A. thaliana* belonging to LAH family (At1g61850, At2G26560, At3G54950, At5G40990, At2G44810, At1G05800, AT1G06800, AT1G53940) and sequences of putative LAH proteins in diatoms (*P. multistriata*, Pmul; *P. multiseriis*, Pmseries; *F. cylindrus*, Fcyl; *P. arenysensis*, Parenys; *P. tricornutum*, Ptri; *S. marinoi*, Smarinoi; *T. pseudonana*, Tpseu. It has been inserted in the alignment also the protein sequence of calcium-independent phospholipase A2 from *H. sapiens* (Hsapiens). Black boxes highlight the hydrolasic motif GxSxG and the catalytic dyad (Serine and Aspartic acid).

3.2 Sequence analysis in order to predict the presence of a putative "*TRANSIT PEPTIDE*" for the proteins that targeting in the chloroplast

The AT1G61850 galactolipase/ phospholipase (A) of *A. thaliana* is localized in chloroplasts. In diatoms, the putative LAH proteins isolated where they are localized? It is possible to predict the presence of a transit peptide and the cellular localization? As it is well known, plastids of diatoms and other chromophytic algae have four surrounding membranes. How the different plastid proteins encoded by the host nuclear genome are transported across all four membranes of diatom plastids? All nucleus-encoded stromal plastid proteins in secondary algae appear to be preceded by an N-terminal extension, which consists of a *signal peptide* and a *transit peptide* (Kroth, 2002).

The characteristic feature of those presequences is the presence of a conserved motif – namely **ASAF** or **AFAP** positioned at the cleavage site of the respective diatom signal peptides. Using bioinformatic tools, the sequences of *P. multistriata*, *P. arenysensis* and *S. marinoi* were analyzed, in order to find the presence of a putative transit peptide. (Figure 26,27,28).The software utilized for the analysis were: *MitoProt* (<http://ihg.gsf.de/ihg/mitoprot.html>), that indicates the presence of a mitochondrial transite peptide, *TargetP*

(<http://www.cbs.dtu.dk/services/TargetP/>), that indicates the presence and length of a transit peptide and the possibility of a subcellular localization, *ChloroP* (<http://www.cbs.dtu.dk/services/ChloroP/>), that indicates the presence and length of a transit peptide and the possibility of a chloroplast localization, *Predotar* v. 1.03 (<http://genoplante-info.infobiogen.fr/predotar/predotar.html>), that indicates the subcellular localization, *PCLR* 0.9 (<http://www.andrewschein.com/cgi-bin/pclr/pclr.cgi>), that predicts the chloroplast localization and *PredSL* (<http://bioinformatics.biol.uoa.gr/PredSL/>), that indicates the presence of a transit peptide and the final destination.

From the analysis of the sequences, only the sequence of *S. marinoi* would seem to possess the characteristics amino acid such as to assume the presence of a signal peptide, for the presence of phenylalanine of the conserved motif, and a cleavage site at position 21-22, then there is not more correspondence suggesting that the signal peptide is lost. In the Figure 29 is shows the amino acid similarity between the protein of LAH of *S. marinoi* and the protein Fcp (fucoxanthin-chlorophyll a/c-binding proteins) of *P. tricornutum* that presents the typical conserved motif.

Software	MitoProt II 1.101	TargetP 1.1	ChloroP 1.1	Predotar v. 1.03	PCLR 0.9	PredSL
Notes	Indicate the presence of a mitochondrial Transit Peptide	Indicate the presence and length of a TP and the possibility of a subcellular localization	Indicate the presence and length of a TP and the possibility of a chloroplast localization	Indicate subcellular localization	Chloroplast Localization Prediction	Indicate the presence of a TP and the final destination
Mitochondria	0.1305 TP= 11aa	0.291		0,03		0.402629
Chloroplasts		0.090	0.465 TP= 31aa	00	0.097	0.001109
Secr.Path/ER		0.018		00		0.001138
Other		0.673		0,97		
Conclusions	low-mitochondrial targeting	Other (SignalP: no cleavage site)	non-Chloroplast Targeting	other	non-Chloroplast Targeting	other

Figure 26: Analysis of protein sequence in *P. multistriata* using the Software MitoProt II, TargetP 1.1, ChloroP 1.1, Predotar v. 1.03, PCLR 0.9, PredSL.

Software	MitoProt II 1.101	TargetP 1.1	ChloroP 1.1	Predotar v. 1.03	PCLR 0.9	PredSL
Notes	Indicate the presence of a mitochondrial Transit Peptide	Indicate the presence and length of a TP and the possibility of a subcellular localization	Indicate the presence and length of a TP and the possibility of a chloroplast localization	Indicate subcellular localization	Chloroplast Localization Prediction	Indicate the presence of a TP and the final destination
Mitochondria	0.2152 TP= no cleavage site	0.072		0.08		0.00067
Chloroplasts		0.05	0.454 TP=37	0.000	0.108	0.048732
Secr.Path/ER		0.392		0.15		0.002438
Other		0.467		0.77		
Conclusions	No mitochondrial targeting	Other (SignalP: no cleavage site)	non-Chloroplast Targeting	other	non-Chloroplast Targeting	other

Figure 27: Analysis of protein sequence in *P. arenysensis* using the Software MitoProt II, TargetP 1.1, ChloroP 1.1, Predotar v. 1.03, PCLR 0.9, PredSL.

Software	MitoProt II 1.101	TargetP 1.1	ChloroP 1.1	Predotar v. 1.03	PCLR 0.9	PredSL
Notes	Indicate the presence of a mitochondrial Transit Peptide	Indicate the presence and length of a TP and the possibility of a subcellular localization	Indicate the presence and length of a TP and the possibility of a chloroplast localization	Indicate subcellular localization	Chloroplast Localization Prediction	Indicate the presence of a TP and the final destination
Mitochondria	0.6265 TP=43aa	0.081		0.01		0.000
Chloroplasts		0.121	0.469 TP=18	0.07	0.032	0.000746
Secr.Path/ER		0.798		0.99		99.999
Other		0.025		0.01		
Conclusions	Possible organelle targeting	Secretory Pathway (SignalP: cleavage site at 21-22 aa)	non-Chloroplast Targeting	ER	non-Chloroplast Targeting	Secretory pathway

Figure 28: Analysis of protein sequence in *S. marinoi* using the Software MitoProt II, TargetP 1.1, ChloroP 1.1, Predotar v. 1.03, PCLR 0.9, PredSL.

<p>P.tricornutum fcpA, fcpB, fcpC, and fcpD genes, complete CDS</p> <p>GenBank: Z24768.1</p> <p>MKFTVFASLFASAAAFAPAQQAARSVATNMAFESELGAQAPLGFYDPLGLVADGDQEKFDRLRYVEIKHGRISMLAVAGYLA QEAGWRLGGDIALDGTKFADI PNGFAALSAIPQAGLIQIIAFIGFLETSMKDI TGGEFVGD FRNGYIDFGWDSFDQETKLRK RAIELNQGRAAQMGILALMVHEQLGVNILPGV</p> <p>...FAPAKNAARTSVATT...</p> <p>LAH <i>Skeletonema marinoi</i></p> <p>MTSHVLVAVCLLLSLPTQCTS↓LRNDAAFTSNYTLNQPSIRKTTVLLYQIEIIDEPLNKR</p>
--

Figure 29: Comparison between the protein of LAH of *S. marinoi* and the protein Fcp (fucoxanthin-chlorophyll a/c-binding proteins) of *P. tricornutum* that presents the typical conserved motif AFAP.

3.3 Validation of internal controls (reference genes) for quantitative gene expression study (qPCR) of *LAH* gene (GENE OF INTEREST *GOI*)

The development of an effective protocol for the study of gene expression using the technique of qPCR can be a valuable tool for the study of the biology of diatoms from the molecular point of view. For these reason I have analysed the expression of a set of genes commonly used as reference genes only throughout the critical phases of the growth curves of cultured *P. multistriata* and *P. arenysensis* diatoms species. The expression of reference gene used for normalization in real-time qPCR analysis should remain constant, in all the individuals studied, under the different experimental conditions considered. In order to determine the best genes that can be used for data normalization in a qPCR experiment, a set of genes was chosen that are usually used as reference genes.

These genes are: histone H4, TBP (TATA box binding protein), involved in basic cellular mechanisms, RPS (30 s ribosomal protein), involved in transcription and protein translation, GAPDH (glyceraldehyde 3-phosphate dehydrogenase), involved in primary metabolism, tubulins (TUB A and TUB B), the constituent protein of microtubules of cells which provide a skeleton for maintaining cell shape, and ACT (actin), a ubiquitous microfilament component, CDKA (cyclin dependent kinase A), involved in the regulation of cell cycle, COPA (coatomer A protein A), involved in

protein transport between the endoplasmic reticulum and Golgi compartments (Ferrante *et al.*, 2009). The putative reference genes analyzed were identified by searching the transcriptomes of *P. multistriata* and *P. arenysensis*. As queries for TBlastN, proteins from the annotated genomes of the diatoms *P. tricornutum* and *T. pseudonana* were used (Siaut *et al.*, 2007). As experimental conditions the main points of a algal growth curve were chosen: exponential phase, stationary phase and declining phase. Samples of *P. multistriata* and *P. arenysensis* were collected during these phases (Figure 30 and 31). Three different cultures of the same strain were set up for each species to obtain biological replicas.

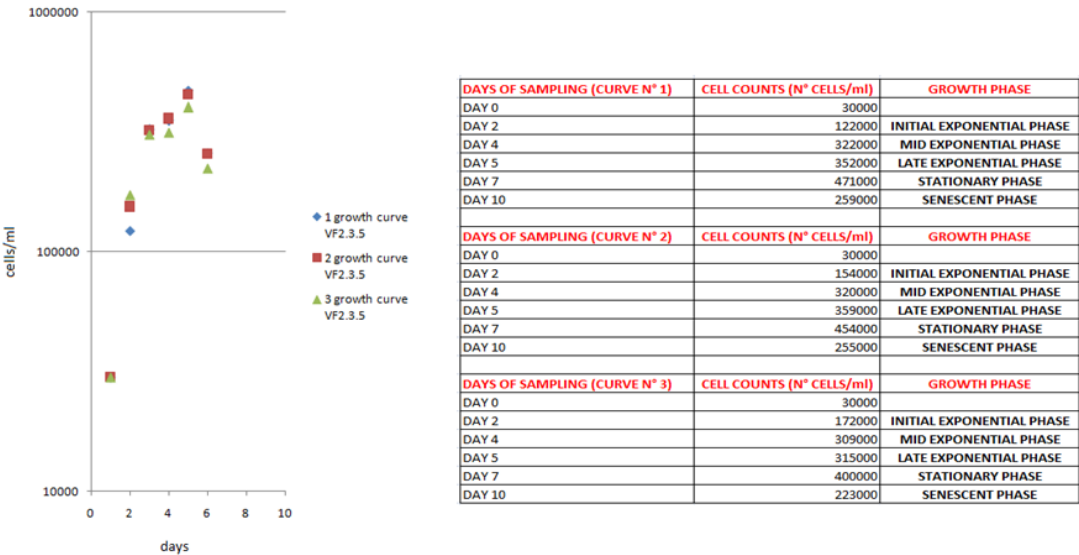
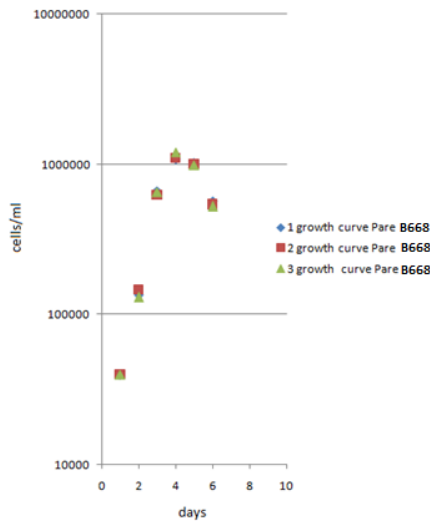


Figure 30: Growth curves of *P. multistriata*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. Days of sampling, cell counts and growth phases are reported in table.



DAYS OF SAMPLING (CURVE N° 1)	CELL COUNTS (N° CELLS/ml)	GROWTH PHASE
DAY 0	40000	
DAY 2	135000	INITIAL EXPONENTIAL PHASE
DAY 4	660000	MID EXPONENTIAL PHASE
DAY 5	1080000	LATE EXPONENTIAL PHASE
DAY 8	1010000	STATIONARY PHASE
DAY 10	567000	SENESCENT PHASE
DAYS OF SAMPLING (CURVE N° 2)	CELL COUNTS (N° CELLS/ml)	GROWTH PHASE
DAY 0	40000	
DAY 2	146000	INITIAL EXPONENTIAL PHASE
DAY 4	623000	MID EXPONENTIAL PHASE
DAY 5	1100000	LATE EXPONENTIAL PHASE
DAY 8	1000000	STATIONARY PHASE
DAY 10	540000	SENESCENT PHASE
DAYS OF SAMPLING (CURVE N° 3)	CELL COUNTS (N° CELLS/ml)	GROWTH PHASE
DAY 0	40000	
DAY 2	130000	INITIAL EXPONENTIAL PHASE
DAY 4	650000	MID EXPONENTIAL PHASE
DAY 5	1200000	LATE EXPONENTIAL PHASE
DAY 8	989000	STATIONARY PHASE
DAY 10	523000	SENESCENT PHASE

Figure 31: Growth curves of *P. arenysensis*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. Days of sampling, cell counts and growth phases are reported in table.

The RNA was extracted at the various points of the growth curve and the all cDNA were tested by RT-PCR using specific primers. The Figure 32 shows the RT-PCR of some cDNA samples of *P. multistriata* and *P. arenysensis* carried out using respectively the primers H4Pm 2FW/H4Pm 2RV that amplify a fragment of 152 bp and H4Pare FW/H4Pare RV which amplify a fragment of 138 bp, and the RT-PCR performed with some of putative reference genes for both species.

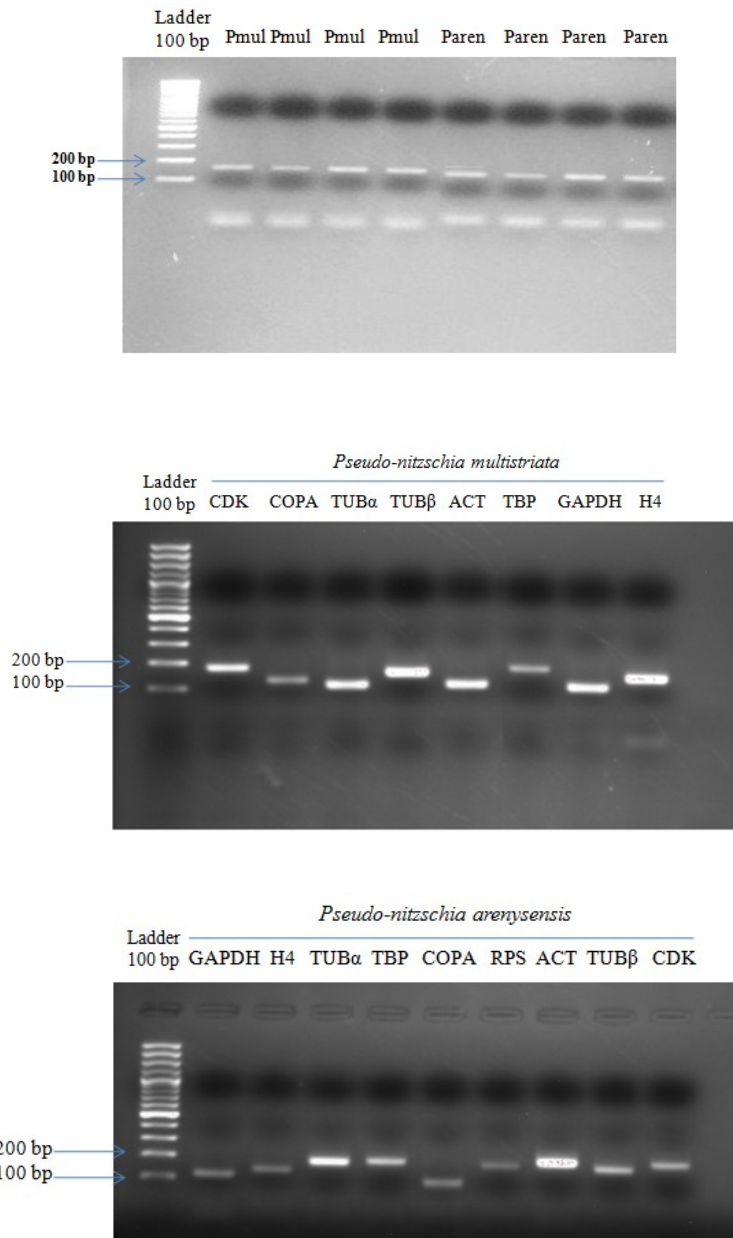


Figure 32: RT-PCR of cDNA samples of *P. multistriata* (Pmul) and *P. arenysensis* (Paren). The amplification fragment for *P. multistriata* was 152 bp, whereas for *P. arenysensis* was 138 bp. RT-PCR of some reference genes in *P. multistriata* and *P. arenysensis*. The sizes of the fragments are between 110 bp and 180 bp. The ladder was the “Ladder 100 bp” (Fermentas).

The primer efficiency was tested according to Pfaffl (2001). For *P. arenysensis* the primers efficiencies of genes calculated from the standard curves were: $E_{H4} = 1,90$, $E_{RPS} = 1,91$, $E_{COPA} = 1,98$, $E_{TUB\ \alpha} = 1,98$, $E_{CDKA} = 1,98$, $E_{TBP} = 2$, $E_{GAPDH} = 2,03$, $E_{TUB\ \beta} = 1,85$. For *P. multistriata* were: $E_{H4} = 2,03$, $E_{RPS} = 2,07$, $E_{COPA} = 1,99$, $E_{TUB\ \alpha} = 1,84$, $E_{CDKA} = 2,04$, $E_{TBP} = 2$, $E_{GAPDH} = 1,79$, $E_{TUB\ \beta} = 1,85$, $E_{ACT} = 1,84$. The final data obtained from the various experiments carried out on the growth curves of *P. multistriata* and *P. arenysensis* were compared using two approaches: *geNorm* and *NormFinder*.

geNorm analysis enables the selection of the optimal set of reference genes from a series of tested candidate reference genes (Vandesompele *et al.*, 2002). The normalization of the genes of interest (GOI) is obtained using the geometric mean of an optimal number of the selected genes. Geometric averaging is used to smooth individual variations in the expression value of a single reference gene; otherwise, large errors could occur in normalized data in samples of interest. This approach needs several candidate genes that reciprocally serve as cross-references to single out the most stable ones. The underlying assumption of this procedure is that ratios between a sample of uniformly expressed, non-normalized HKGs should remain relatively constant among a suitable selection of HKGs. To show this, M (stability value) values are calculated, corresponding to the average pairwise variation of a single HKG candidate to all other genes. Low M values indicate stable

expression. In iterative steps, genes with the lowest stability and consequently with the highest M value, are removed. A new M value for each remaining gene is calculated until only 2 genes remain. For the analyses, Ct values were converted into relative quantities for analysis with geNorm, considering the PCR efficiencies of the genes (Vandesompele *et al.*, 2002; Meller *et al.*, 2005).

NormFinder is a mathematical approach for identifying the optimal gene to use as reference among a set of candidates. It ranks the set of candidate reference genes according to their expression stability in a given sample set and given experimental design. The algorithm is rooted in a mathematical model of gene expression and uses a solid statistical framework to estimate not only the overall expression variation of the candidate normalization genes, but also the variation between sample subgroups of the sample set. Notably, “NormFinder” provides a stability value for each gene, which is a direct measure for the estimated expression variation enabling the user to evaluate the systematic error introduced when using the gene for normalization; this approach is based on a two-way ANOVA. Regarding the analysis of the expression of genes along the growth curves of *P. multistriata*, putative reference genes analyzed show different levels of stability. The threshold of suitability for HKG’S M values should be < 1.5 (Strube *et al.*, 2008; Van Hiel *et al.*, 2009), and none of the analyzed genes shows a M

value $> 1,5$, indicating that they could all be used as control genes. However there is a degree of variability, the results obtained from the analysis carried out with *geNorm* (Figure 33) show that the gene with lowest stability and consequently with the highest value of M is GAPDH, while the genes with highest stability and lowest value of M are represented by CDK, TUB A and ACT. The analysis carried out with *NormFinder* (Figure 34) shows that TUB B is the best candidate as a reference gene, indeed it shows the smallest change of expression along the growth curves.

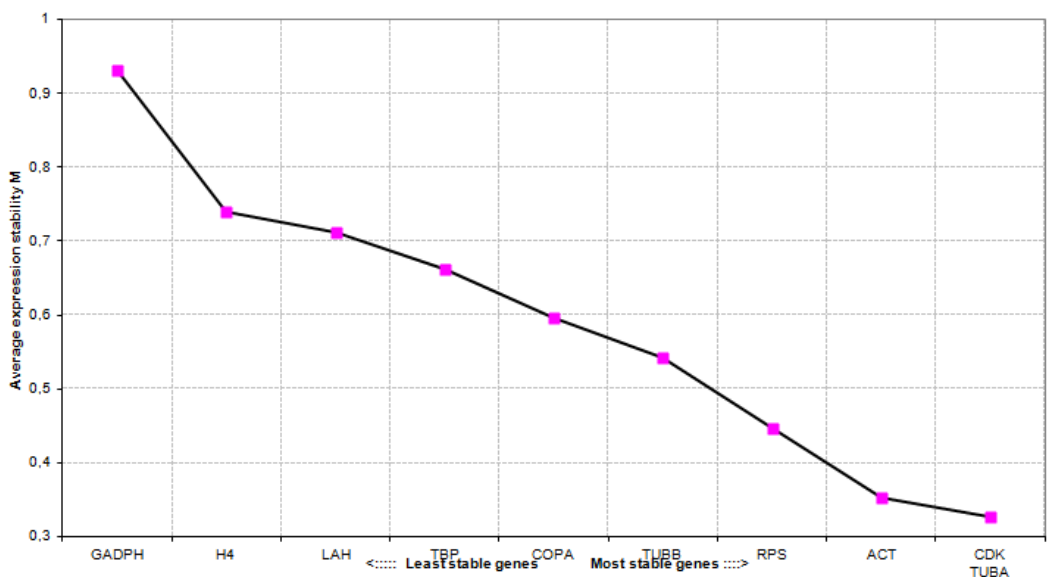


Figure 33: *geNorm*- based analysis on average values from three replicas of samples collected along the growth curve of *P. multistriata*.

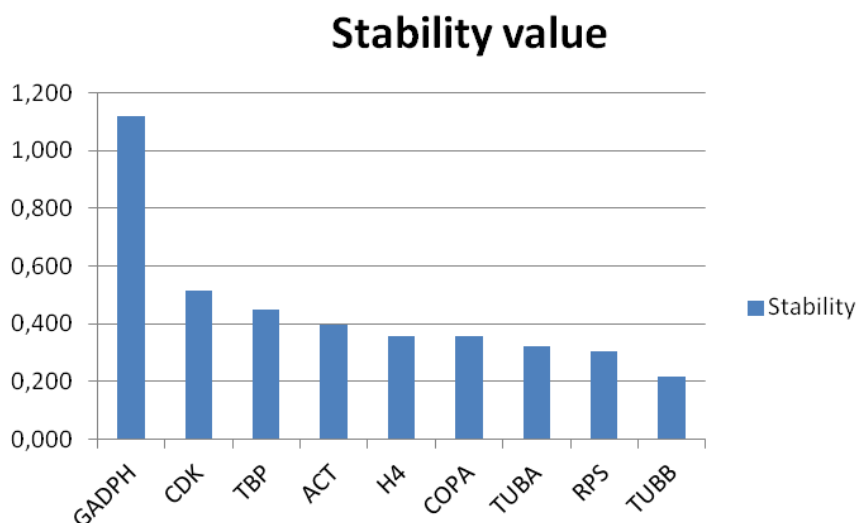


Figure 34: NormFinder- based analysis on average values from three replicas of samples collected along the growth curve of *P. multistriata*.

Regarding the analysis of the expression of genes during the growth curves of *P. arenysensis* the geNorm-based analysis on media of three growth curves shows that the gene with lowest stability and consequently with the highest value of M is COPA(Figure 35), but also in this case none of the analyzed genes shows a M value >1,5. The analysis carried out with *NormFinder* (Figure 36) shows that CDK is the best candidate as a reference gene, indeed it shows the smallest change of expression along the growth curves.

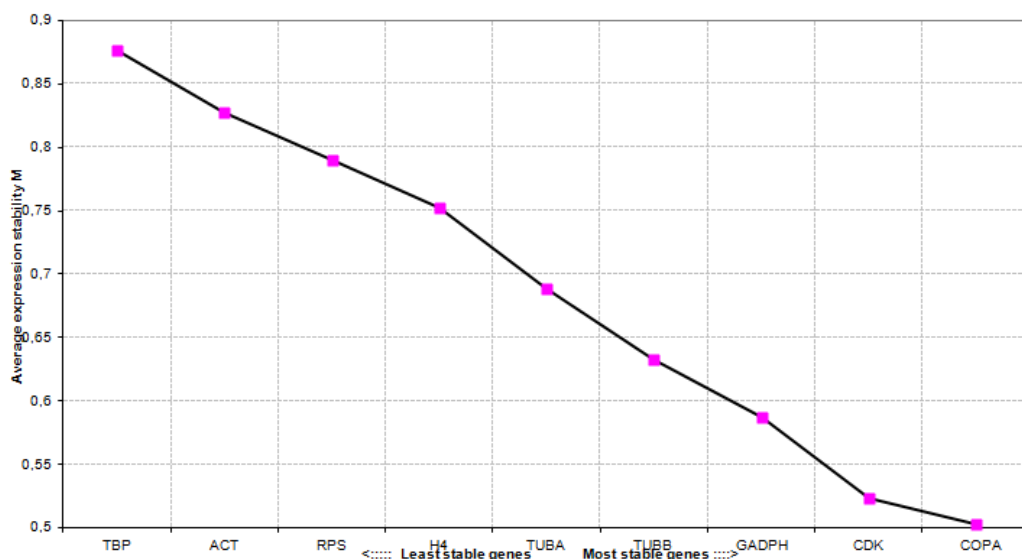


Figure 35: geNorm- based analysis on average values from three replicas of samples collected along the growth curve of *P. arenysensis*.

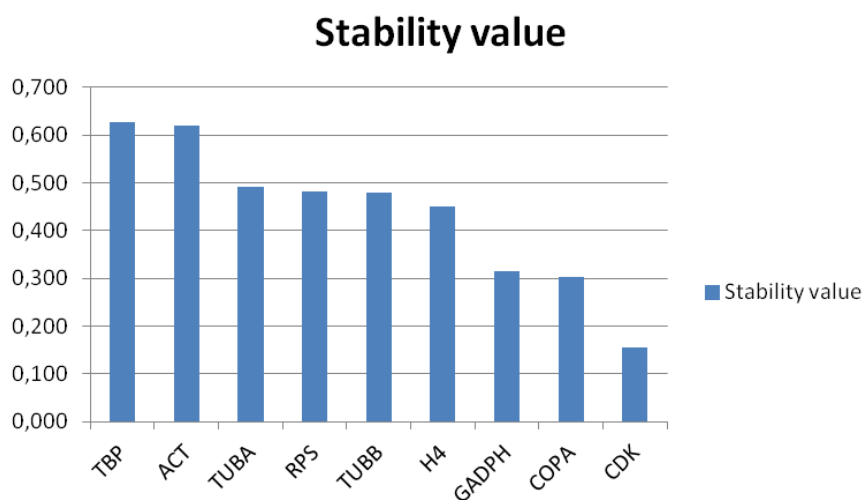


Figure 36: NormFinder- based analysis on on average values from three replicas of samples collected along the growth curve of *P. arenysensis*.

3.4 Gene expression analysis of putative *LAH* gene in *P. multistriata*

Oxylipins synthesis varies during the phases of the cycle of algal growth in a manner dependent on the density of population. It was clear that there is a modulation in the biosynthesis of these compounds at biochemical level and that a hydrolytic activity, related to LAHs, is involved in the production of oxylipins. In order to verify if this variation at physiological level was the effect of a modulation at the transcriptional level, the relative expression level of putative gene encoding LAH protein was analyzed using the qPCR technique. The expression profile of the LAH gene in *P. multistriata* was analyzed. Starting from the consideration that the oxylipin compounds may represent a chemical signal of end of bloom, three type of experiments were set up starting from a different number of cells inoculated. In the first two types of experiments only the final stages of growth were analyzed (stationary and declining phase), in order to see an evident difference in gene expression between the beginning of the growth and the final stages, while in the third, all main stages of growth were analyzed, to validate a possible modulation in gene expression along the entire growth curve.

Experiment 1: At day 0 18,000 cells/ml were inoculated and samples were collected from four points along the growth curve: mid exponential phase (day 3 from inoculum), late exponential phase (day 5), stationary phase (day 7), declining phase (from day 10 onwards). The growth was monitored

by cell counting and observation with an optical microscope. The curves performed in duplicate show a similar trend during the various growth phases, even if the number of cells is slightly different from one curve to another (Figure 37). The samples were collected during the various days always the same time, to minimize alterations in the levels of expression of the gene of interest and of the genes used as a “reference genes”. The RNA was extracted at the various points of the growth curve and the cDNA was tested by RT-PCR using specific primers. Preliminary qPCR experiments were performed to determine the optimal concentration of cDNA, the primers efficiency and the suitable set of genes that could be used as reference genes, in order to normalize the data. The primers used for qPCR experiments (LAHPm 1FW/LAHPm 6RV) amplify a fragment of 167 bp. The normalization of the expression profile of LAH gene was carried out using the $\Delta\Delta C_t$ method (Pfaffl, 2001), a method described in Materials and Methods. Arbitrarily as a “reference condition” for the analysis of the results obtained, it was considered the early exponential phase, since it was assumed that, if the modulation in the synthesis of oxylipins was due to a transcriptional event, in the initial stages of the growth curve the expression of the gene coding for the LAH protein must be basal as if it were a condition *wild type*. The expression profile of the LAH gene was normalized with H4 and RPS genes. As shown in the Figure 37, the expression of RPS and H4 has

remained fairly constant among the different phases of growth, while the profile of expression of LAH gene varies slightly randomly in the various conditions analyzed. Considering the fold changes in gene expression (cut off > 2), these vary by about 0.6-fold during the late exponential phase, by about 0.7-fold during the stationary phase, and by about 1-fold during the declining phase. The results reported in the graph represent the average of the values obtained by the two growth curves. The values of ΔCt (the difference in threshold cycles for target gene and reference) are: $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{RPS}) = 3,27$; $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{H4}) = 3,61$.

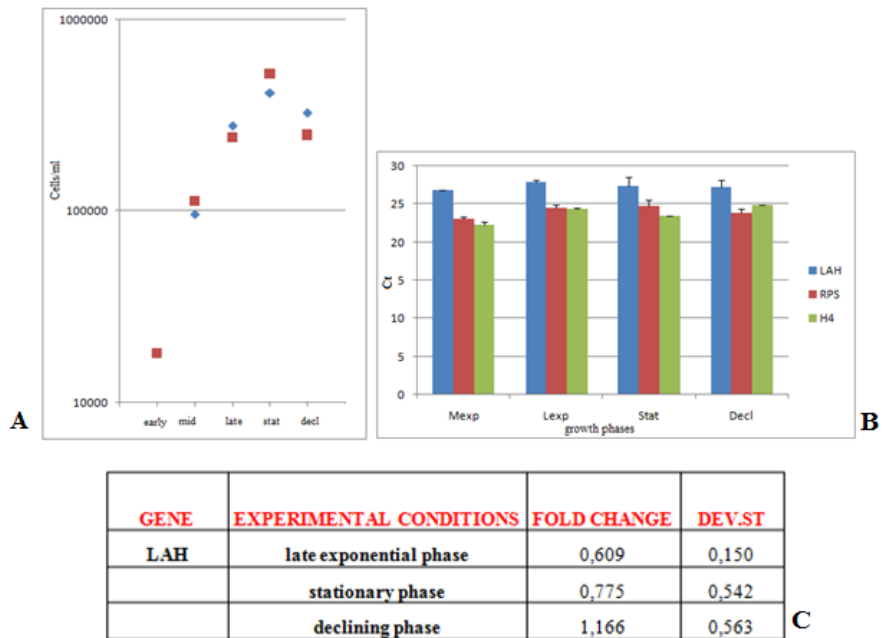


Figure 37: A) Growth curves of *P. multistriata*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale B) qPCR cycle threshold values of the LAH gene (target gene), RPS gene and H4 gene (reference genes) in samples along the curves. C) Expression fold change of LAH gene during the experimental conditions analyzed.

Experiment 2: At day 0 36,000 cells / ml were inoculated, and the samples were collected from four points along the growth curve: early exponential phase (day 2 from inoculum), late exponential phase (day 7), stationary phase (day 8), declining phase (from day 9 onwards). Although the trend of the curves is similar, the two curves differ slightly in number of cells, which is particularly evident in the final phases, where in the second curve the culture has reached in stationary phase a number of cells equal to 1×10^6 cells/ml. The results reported in the graph represent the average of the values obtained by the two growth curves. The expression levels of LAH gene were compared with those of H4 and RPS genes. The expression levels of reference genes remained constant during the growth phase, while the expression levels of LAH gene, as in the previous experiment, show only a slightly variation randomly along the growth curve. Considering the fold changes in gene expression, the variation was by about 1,7-fold during the late exponential phase, by about 1,3-fold during the stationary phase, and by about 1,2-fold during the declining phase (Figure 38). The values of ΔCt are: $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{RPS}) = 4,54$; $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{H4}) = 4,33$.

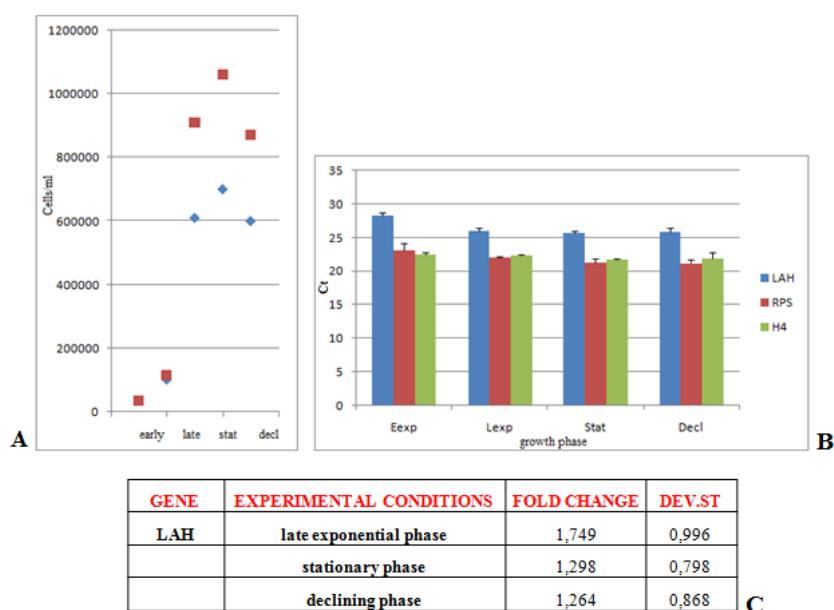


Figure 38: A) Growth curves of *P. multistitata*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. B) qPCR cycle threshold values of LAH gene (target gene), RPS gene and H4 gene (reference genes) in samples along the curves. C) Expression fold change of LAH gene during the experimental conditions analyzed.

Experiment 3: the third experiment was carried out by sampling several days along the growth curve. Three growth curves were followed in triplicate, however they did not show the same trend, in fact, while the first and the second curves have reached the declining phase after 10 days, the third has a long stationary phase still maintained after 10 days. At day 0 30,000 cells/ml were inoculated, and the samples were collected from five points along the growth curve based on the cell counts. For the first growth curve: early exponential phase (day 2 from inoculum), mid exponential phase

(from day 4 to 5), late exponential phase (day 6), stationary phase (from day 7 to 9), declining phase (day 10). For the second growth curve: early exponential phase (day 2 from inoculum), mid exponential phase (day 4), late exponential phase (day 5), stationary phase (from day 6 to 9), declining phase (day 10). For the third growth curve: early exponential phase (from day 2 to 4), mid exponential phase (from day 5 to 6), late exponential phase (day 7), stationary phase (from day 8 to 10). The results reported in the graph represent the average of the values obtained by the three growth curves. The expression levels of LAH gene along the growth curves were compared to those of RPS and H4 genes. Also in this experiment analysing several days along the growth curve, the expression levels of RPS and H4 remained mostly constant during the various phases, while, the slightly variations observed for the gene LAH are not significant nor would seem to be a clear modulation in the expression levels. Considering the fold changes in gene expression, the variation was by about -0,26-fold during the mid exponential phase, 0,57-fold during the late exponential phase, by about -0,36-fold during the stationary phase, and by about -0,62-fold during the declining phase (Figure 39). The values of ΔCt are: $\Delta Ct (Avg. Ct_{LAH} - Avg. Ct_{RPS}) = 0,88$; $\Delta Ct (Avg. Ct_{LAH} - Avg. Ct_{H4}) = 2,18$.

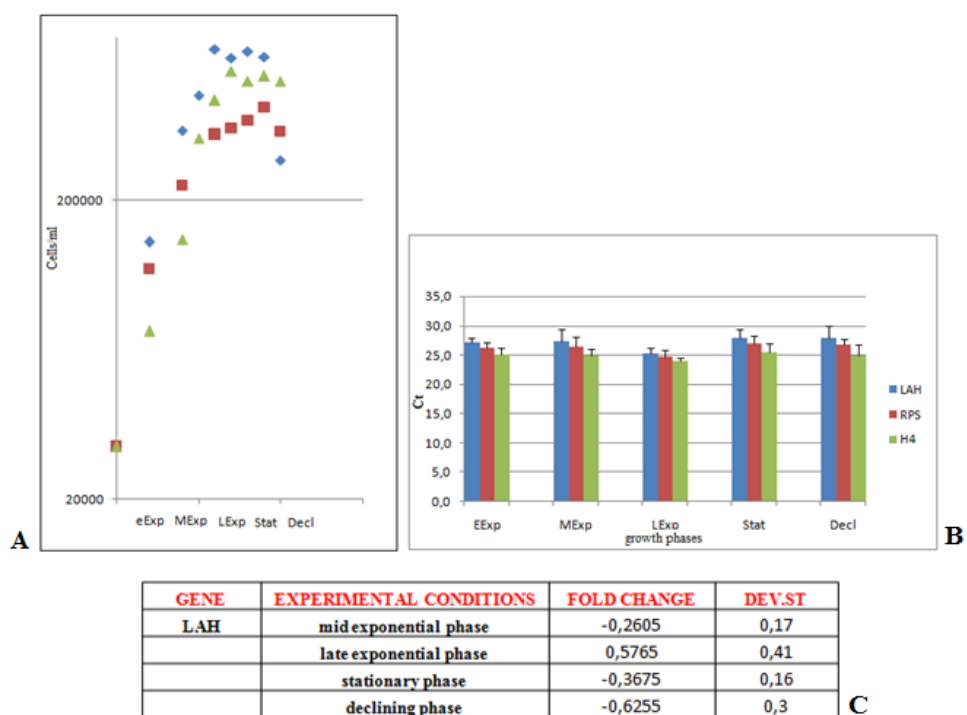


Figure 39: A) Growth curves of *P. multistriata*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. B) qPCR cycle threshold values of the LAH gene (target gene), RPS gene and H4 gene (reference genes) in samples along the curves. C) Expression fold change of LAH gene during the experimental conditions analyzed.

3.5 Gene expression analysis of putative *LAH* gene in *P. arenysensis*

The same criteria that have guided the experiments in *P. multistriata*, were applied to those carried out in *P. arenysensis*. The relative expression level of putative gene encoding LAH protein was analyzed using the qPCR techniques. The preliminary experiments about the suitable set of genes that could be used as reference genes in a qPCR analysis have shown that genes more stable for *P. arenysensis* are COPA (coatomer A protein A) and CDKA (cyclin dependent kinase A) genes. The primers used for qPCR experiment (LAHPare FW/LAHPare RV) amplify a fragment of 179 bp. Three algal growths were performed. The biological replicas, as shown in the Figure, were perfectly comparable. At day 0 40,000 cells/ml were inoculated, and the samples were collected from five points along the growth curve based on the differences provided by the three cultures: early exponential phase (day 2 from inoculum), mid exponential phase (day 4), late exponential phase (day 5), stationary phase (day 8), declining phase (day 10). The results reported in the graph represent the average of the values obtained by the three growth curves. Like *P. multistriata*, the expression levels of LAH gene were not significant, compared to those of reference genes COPA and CDK. The fold changes in gene expression were by about -2,7-fold during the mid exponential phase, -2,5-fold during the late exponential phase, by about -3-fold during the stationary phase, and by about -2,3-fold during the declining

phase (Figure 40). The values of ΔCt are: $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{COPA}) = -0,45$; $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{CDK}) = -2,83$.

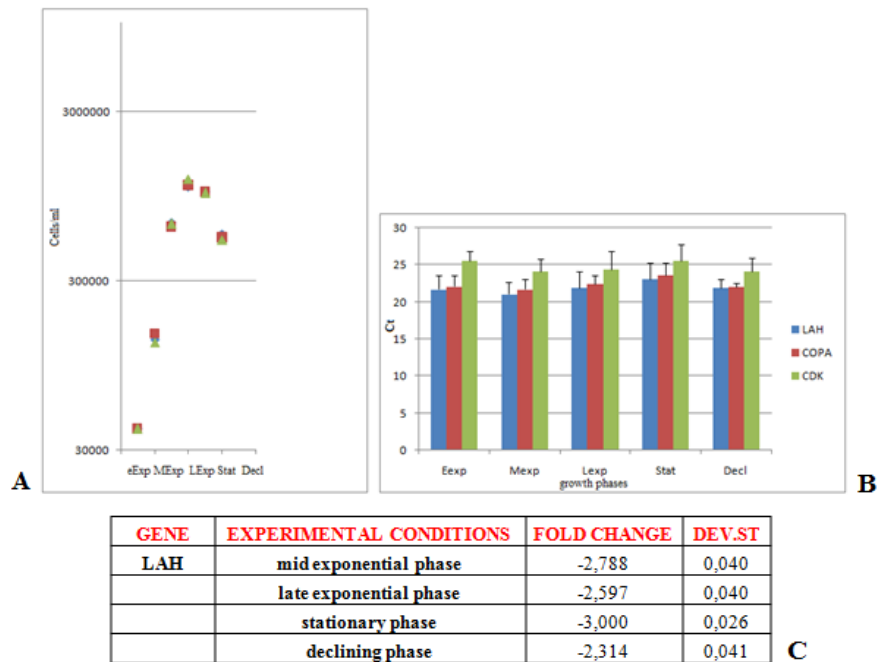


Figure 40: A) Growth curves of *P. arenysensis*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. B) qPCR cycle threshold values of LAH gene (target gene), COPA gene and CDK gene (reference genes) in samples along the curves. C) Expression fold change of LAH gene during the experimental conditions analyzed.

3.6 Gene expression analysis of putative *LAH* gene in *S. marinoi*

The search in the transcriptome of *S. marinoi* led to the isolation of the putative LAH gene. Preliminary experiments on the validation of genes that could to be used as reference genes in the case of *S. marinoi* have not been performed in this work of thesis (data access is made possible by courtesy of Dr Orefice and Dr Romano at the Zoological Station “Anton Dohrn”). On the basis of the results obtained, the genes that could be used as reference genes are the actin (ACT) and ubiquitin (UBI). The expression profile of LAH genes was compared to those of two genes. qPCR experiments were conducted on three algal growths of *S. marinoi*. The cultures showed the same trend during the sampling period. At day 0 5000 cells / ml were inoculated, and the samples were collected from five points along the growth curve: early exponential phase (from day 3 to day 4), mid exponential phase (from day 5 to day 6), late exponential phase (day 7), stationary phase (day8), declining phase (from day 9 to 10). The primers used for qPCR experiments (LAHSke FW/LAHSke RV) amplify a fragment of 186 bp. The results reported in the graph represent the average of the values obtained by the three growth curves. As show in Figure 41 there would not seem to be a modulation in the levels of expression of the LAH gene along the growth curve, despite a slight upward trend compared to those of ACT and UBI genes is visible (considering the cycle threshold values Ct), whose expression

levels appear to be fairly homogeneous. Infact, considering the fold changes in gene expression, the variation was by about 0,01-fold during the mid exponential phase, -0,28-fold during the late exponential phase, by about -0,93-fold during the stationary phase, and by about -1,25-fold during the declining phase. The values of ΔCt are: $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{ACT}) = -1,9$; $\Delta Ct (\text{Avg. Ct}_{LAH} - \text{Avg. Ct}_{UBI}) = -2,1$.

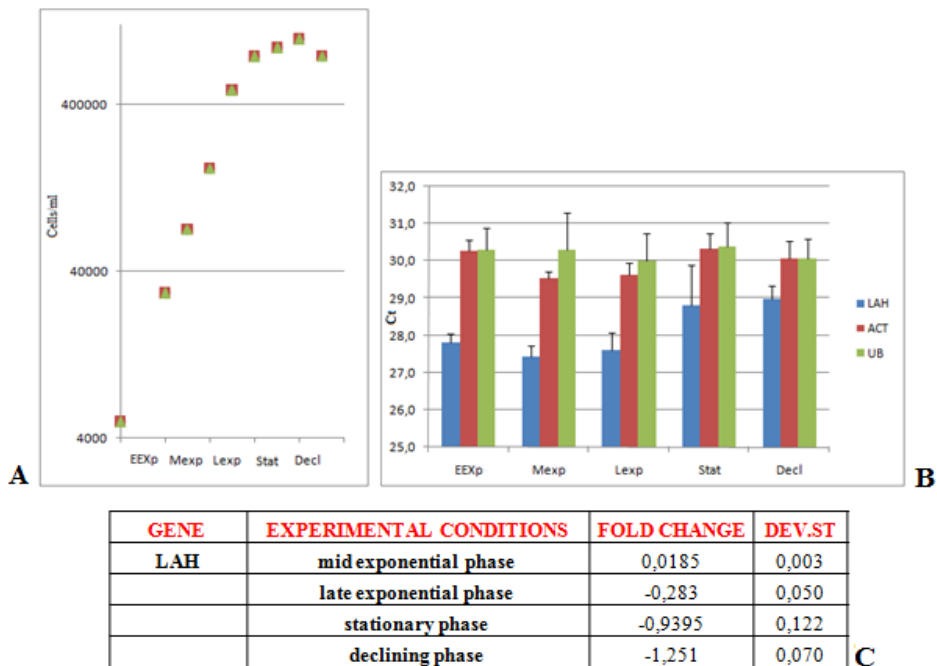


Figure 41: A) Growth curves of *S. marinoi*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale. B) qPCR cycle threshold values of LAH gene (target gene), actin (ACT) gene and ubiquitin (UBI) gene (reference genes) in samples along the curves. C) Fold change of LAH gene during the experimental conditions analyzed

3.7 Characterization of LOX pathway and oxylipin profiling in *P. multistriata*

In addition to the analysis of gene expression, a study of oxylipin metabolism was performed by Liquid Chromatography-Mass Spectrometry (LC-MS) in positive ESI⁺ mode, in agreement with Cutignano *et al.*, 2011. This analytical technique is employed for post-genomic studies based on metabolite profiling and couples chromatographic resolution of complex mixtures of organic compounds with mass or mass/mass (MS/MS) analysis of each component. In order to minimize any variable technique in comparing molecular and chemical data, LC-MS analyses were carried out in triplicate on the same samples that were used for qPCR. For each diatom sample, cells were collected along the growth in order to have representative samples of early exponential phase, mid exponential phase, late exponential phase, stationary phase, and declining phase (Figure 42).

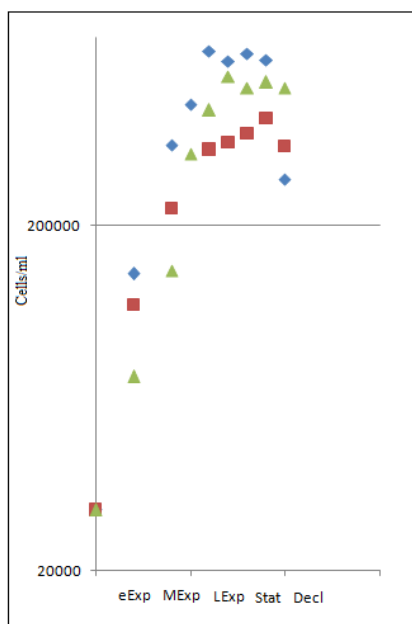


Figure 42: Growth curves of *P. multistriata*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale.

The LC-MS profile of oxylipin extracts of *P. multistriata* showed two major peaks with m/z 371f ($M+Na^+$) and elution at 15.8 and 17.5 min, as well as two peaks with m/z 355, retention time (R_t) of 20.7 and 22.5 and UV absorption centered at 234 nm for both compounds. The first two peaks were identified as epoxyalcohols of eicosapentaenoic acid on the basis of MS/MS fragmentation pattern generated by high collision energy (Cutignano *et al.*, 2011). In particular, the peak at 15.8 min revealed a single fragment at m/z 273 derived from 16-hydroxy-14,15-epoxy-5Z,8Z,11Z,17Z-eicosatetraenoic acid (16,14 HepETE) (Figure 43). The other epoxyalcohol (R_t = 17.5 min) showed a single fragment at m/z 153, that was consistent with the occurrence

of 7-hydroxy-5,6-epoxy-8Z,11Z,14Z,17Z- eicosatetraenoic acid (7,5 HepETE) (Figure 44).

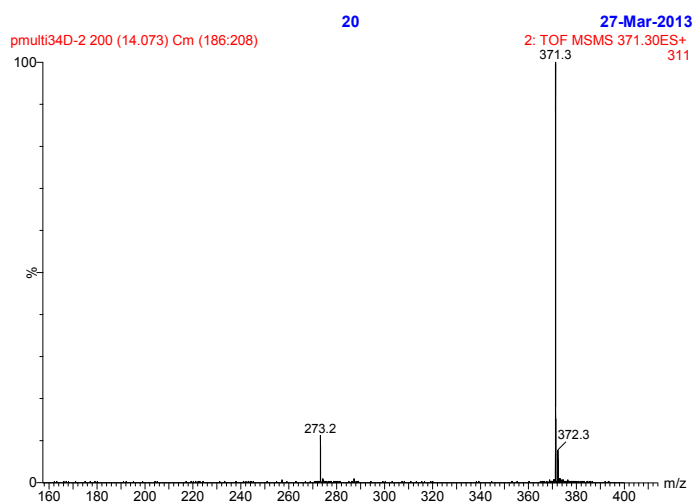


Figure 43: MS/MS spectrum of 16-hydroxy-14,15-epoxy-5Z,8Z,11Z,17Z-eicosatetraenoic acid (16,14 HepETE) from the extract of *P.multistriata*. Numbers above peaks indicate the m/z value of the pseudomolecular ion (371, for $M+Na^+$) and the diagnostic fragment.

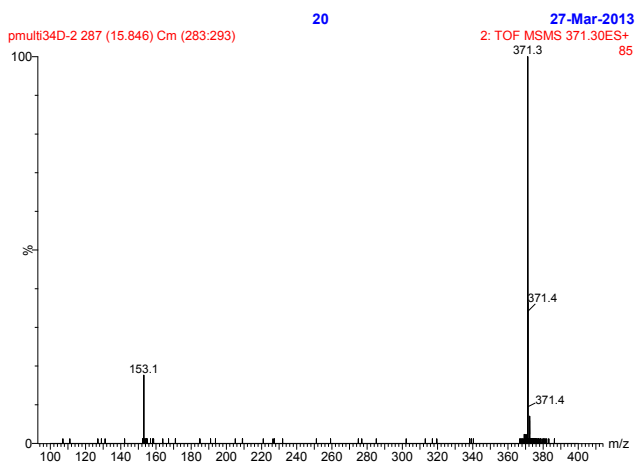


Figure 44: MS/MS spectrum of 7-hydroxy-5,6-epoxy-8Z,11Z,14Z,17Z- eicosatetraenoic acid (7,5 HepETE) from the extract of *P.multistriata*. Numbers above peaks indicate the m/z value of the pseudomolecular ion (371, for $M+Na^+$) and the diagnostic fragment.

The MS/MS data established the position of oxygenation of EPA in agreement with operation of 14-LOX and 5-LOX in *P. multistriata* (Figure 45). The corresponding hydroxyacids were assigned on the basis of retention time on reverse phase HPLC as 14-hydroxy-5Z,8Z,11Z,15E,17Z-eicosapentaenoic acid (14-HEPE) and 5-hydroxy-6E,8Z,11Z,14Z,17Z-eicosapentaenoic acid (5-HEPE).

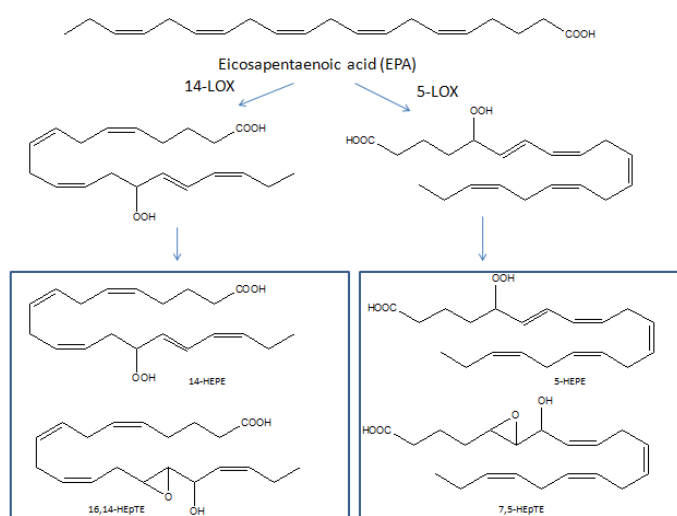


Figure 45: Lipoxigenase pathways operating in *P. multistriata*. For convenience, EPA is reported as free fatty acids.

Measurements of oxylipin along the growth curve indicated persistent occurrence of LOX products (Figure 46), even if variation of these products suggested a constant level for 5-LOX activity and modulation of 14-LOX. Quantitatively, 14-LOX products were largely predominant mainly and 14-HEPE/16,14-HEpTE were always detectable in the LC-MS profile of *P.*

multistriata, even if their concentration progressively increased before stationary phase. 5-HEPE was rather constant along the growth, whereas 7,5-HEpTE was evident only in the declining phase.

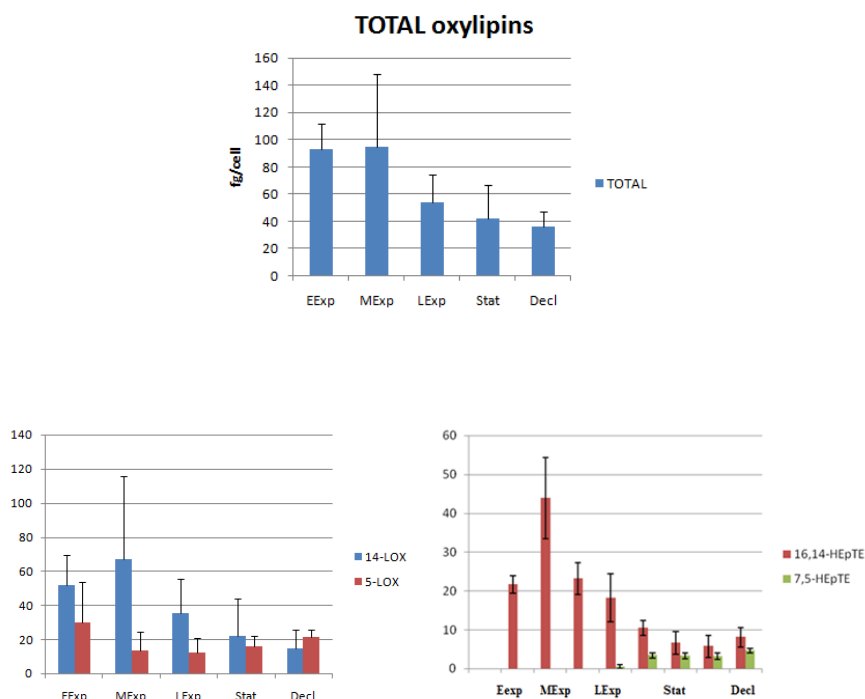


Figure 46: LC-MS profiling of oxylipins (fg/ml) along the growth curves of *P. multistriata*.

3.8 Characterization of LOX pathway and oxylipin profiling in *P. arenysensis*

The LCMS profile of oxylipin extracts of *P. arenysensis* showed two major peaks in HPLC chromatogram for 13-hydroxy-14,15-epoxyeicosa-5Z,8Z,11Z,17Z-tetraenoic acid (13,14-HEpETE) (m/z 371, $M+Na^+$) at 14.29, λ_{max} 210) and 14-hydroxy-5Z,8Z,11Z,13E,17Z-eicosapentaenoic acid (15-HEPE), which is in agreement with 15S-LOX metabolism (Figure 47) previously described by d'Ippolito *et al.* (2009) in this species (reported as *P. delicatissima*). However it is noteworthy that the strains also produced minor

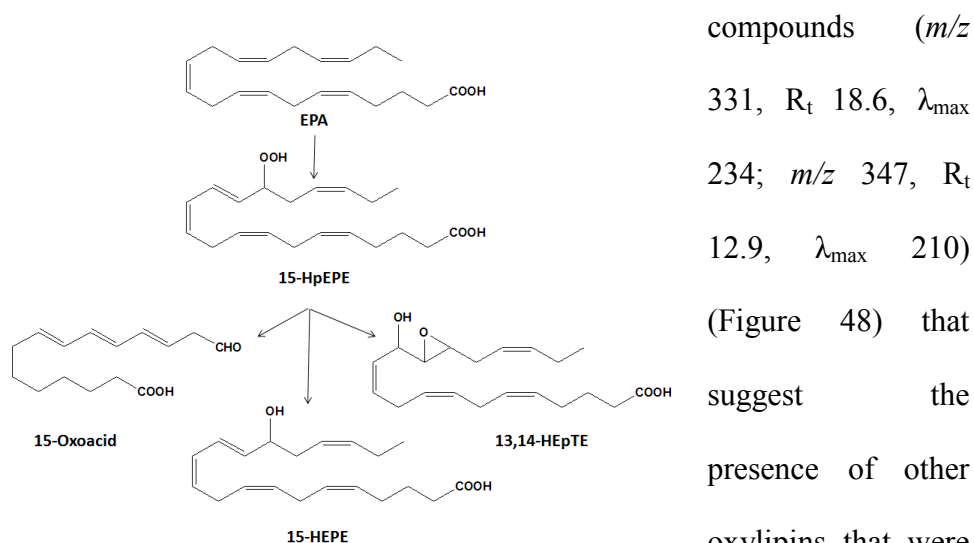


Figure 47: 15S-LOX pathway in *P. arenysensis*.

studied (Giuliana d'Ippolito, personal communication).

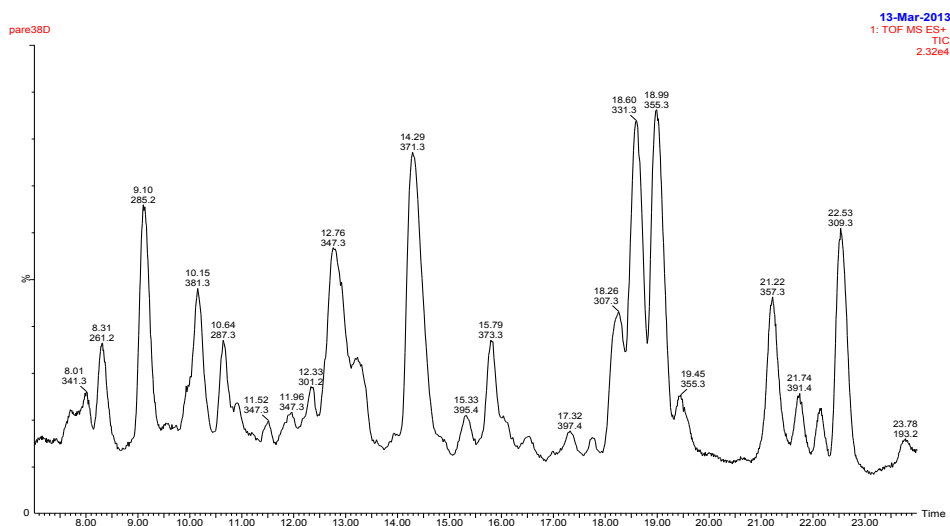


Figure 48: LC/MS profile *P. arenysensis*. Numbers above peaks indicate the molecular ion (M+Na⁺) of the more abundant compounds, determined by ESI⁺ ionization.

Analysis of LOX products was carried out on three individual cultures of the

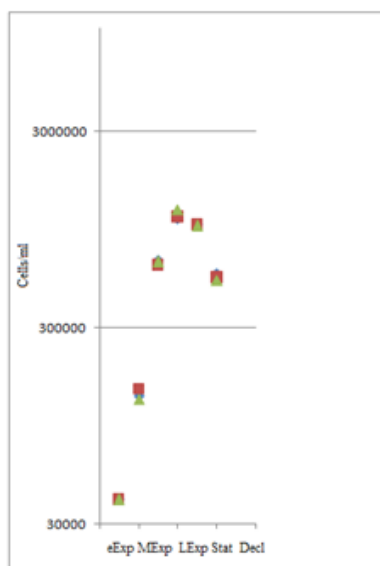


Figure 49: Growth curves of *P. arenysensis*: the graph shows the rate of growth (number of cells/ml) plotted on a log scale.

microalga with samples that were also used for RNA qPCR at early exponential phase (day 2 from inoculum), mid exponential phase (day 4), late

exponential phase (day 5), stationary phase (day 8), declining phase (day 10) (Figure 49).

Oxylipin profiling by LC-MS gave results similar to those above described for *P. multistriata*, since total quantity of oxylipins/cell progressively increased during the early and mid exponential phase and reached the highest values (160 fg/cell) in late exponential. During the last part of the growth phases (late, stationary and declining phase) the level of oxylipins decreased to 30-80 fg/cell.

In agreement with previous studies on 15S-LOX metabolism in *P. arenysensis* (Figure 48), the oxylipin mixture of this species was composed of 15S-hydroxy-(5Z,8Z,11Z,13E,17Z)-eicosapentaenoic acid (15S-HEPE), 15-oxo-5Z,9E,11E,13E-pentadecatetraenoic acid and 13,14-*threo*-13R-hydroxy-14S,15S-*trans*-epoxyeicosa-5Z,8Z,11Z,17Z-tetraenoic acid (13,14-HEpETE). The synthesis of 15-oxoacid is detectable from the late exponential phase until stationary phase and tend to decrease during the declining phase.

3.9 Characterization of LOX pathway and oxylipin profiling in *S. marinoi*

The characteristic profile of *S.marinoi* showed two major groups of products in agreement with synthesis from EPA and C16-polyunsaturated fatty acids, namely hexadeca-12Z,9Z,6Z-trienoic acid (HDTA) and hexadeca-15,12Z,9Z,6Z-tetraenoic acid (HDTrA) (Fontana *et al.*, 200; d'Ippolito *et al.*, 2005). As shown in Figure 50, the chromatographic profile contained two major epoxyalcohols with m/z 371 ($M+Na^+$) at 14,1 min and m/z 319 ($M+Na^+$) at 8,3 min, consistent with the structures of 13-hydroxy-14,15-epoxyeicosa- 5,8,11,17-tetraenoic acid (13,14-HepETE) and 11-hydroxy-9,10-epoxyhexadeca- 6,12-dienoic acid (11,9-HepHDE), respectively.

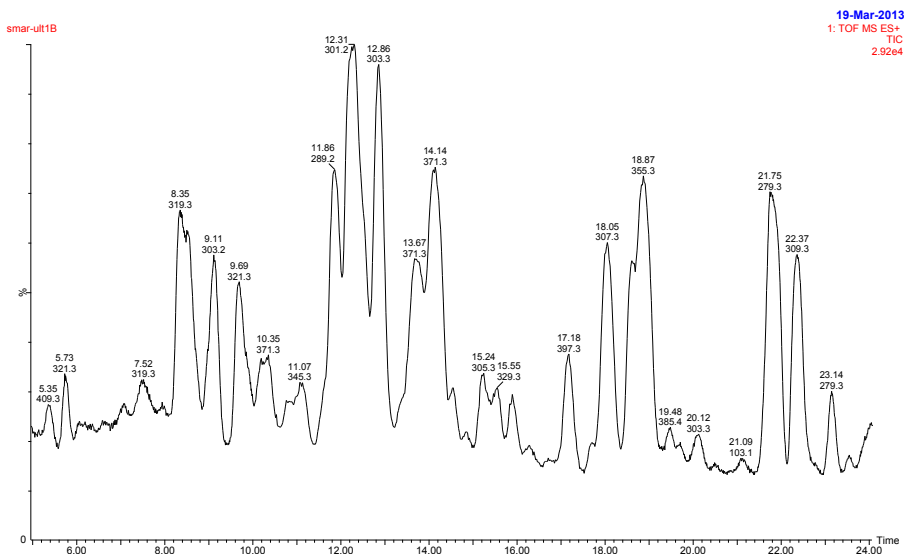


Figure 50: LC/MS profile *S. marinoi*. Numbers above peaks indicate the molecular ion ($M+Na^+$) of the more abundant compounds, determined by ESI+ ionization.

MS/MS fragmentation fully confirmed these assignments thus establishing the presence of at least EPA-dependent 15S-LOX pathway (Fontana *et al.*, 2007) together with the unusual 9-LOX activity dependent by HDTA and HDTrA (d'Ippolito *et al.*, 2005) (Figure 51).

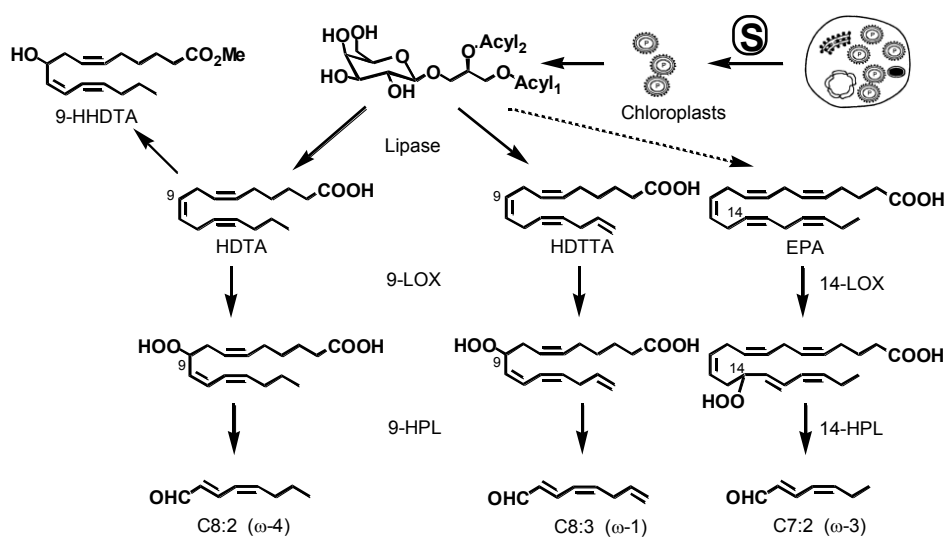


Figure 51: LOX metabolism in *S. marinoi*.

4. DISCUSSION

Prokaryotes and eukaryotes produce a wide range of organic compounds, most of which have specific functions that vary from species to species. These substances, traditionally known as “*secondary metabolites*” or “*natural products*”, are product of specific, secondary pathways often distributed among limited taxonomic groups. Usually these compounds also possess specific functions in the adaptive protection against predators and microbial infections, or in inter- or intra-specific signals, such as attractants for pollinating insects. Within marine planktic communities, an important topic is the potential role of secondary metabolites as defensive or allelopathic agents (allelochemicals affect competition between plant species) in controlling species successions during bloom development, species competition and defense against zooplanktic grazers, mainly copepods (Ianora *et al.*, 2009).

Chemical interactions are very well known and studied in terrestrial ecosystems (Inderjit & Duke, 2004), but studies in aquatic systems have been biased by technical difficulties, mainly arising from dilution in the water medium and physical constraints such as viscosity or shear forces (Wolfe, 2000). In aquatic systems there is a broader diversity of species and chemical compounds than in terrestrial ecosystems (McClintock *et al.*, 2001).

In the recent years, attention of some research groups has been focusing on the eco-physiological role of diatoms *oxylipins*, oxygenated derivatives of fatty acids, which are involved in defense and adaptation induced by abiotic and biotic stimuli. In plants the biosynthesis of oxylipins occurs in response to environmental stress and developmental cues via a pathway that involves the lipase-mediated release of polyunsaturated fatty acids (PUFAs) from membrane lipids. Linoleic and linolenic acid are the major polyunsaturated fatty acids in plant tissues and for these reason the most common substrates for LOXs, leading to two possible products, the 9- and 13-hydroperoxy fatty acids. In diatoms, oxylipins have been particularly studied for the antimitotic activities and the consequent block of embryogenesis and induction of teratogenic effects in copepods, small crustaceans herbivores present both in marine and freshwater habitats. Pohnert showed (2002) that activation of oxylipin-based chemical defense in the diatom *T. rotula* is initiated by phospholipases that act immediately after cell damage, and that the main enzyme activity responsible for initiation of the aldehyde-generating lipase/lipoxygenase/hydroperoxide lyase cascade was characterized as a phospholipase A2. In 2004, d'Ippolito *et al* described for the first time that the biochemical pathways leading to volatile aldehydes from complex lipids in the diatom *S. costatum*, in analogy with plants operates through hydrolysis of glycolipids (GL) and release of EPA and C16

PUFAs, which feed the downstream lipoxygenase (LOX)/hydroperoxide lyase (HPL) pathway. These authors also suggested that the process could be triggered by a novel galactolipid hydrolyzing activity that, like lipolytic acyl hydrolases (LAHs) of plants, is able to release fatty acids from phospholipids and glycolipids. A similar mechanism has been more recently demonstrated in *T. rotula* (Cutignano *et al.*, 2007). In this centric diatom, formation of free fatty acids and oxylipins require hydrolysis of polar lipids but not triglycerides. Preliminary characterization of the lipolytic enzyme yielded an active protein fraction of putative weight of 40–45 kDa that was able to produce fatty acids and *lyso* derivatives of galactolipids, phospholipids but not triglycerides, as much as is currently known for plant LAHs (Cutignano *et al.*, 2007). The enzymatic analogies from taxonomically unrelated diatom species led also the authors to put forward a general model for producing high local concentrations of ecologically relevant substances by the exploitation of cellular resources as abundant as chloroplastic glycolipids, with little or no additional energy costs during normal growth of healthy individuals. Synthesis of oxylipins in *T. rotula* (d'Ippolito *et al.*, 2003 and 2005) and *S. costatum* (now revised to *S. marinoi*) (Fontana *et al.*, 2007) points out the role of chloroplasts in providing signaling molecules that mediate eco–physiological processes of marine diatoms.

According to these considerations, the main aim of this doctoral thesis was to provide molecular tools to investigate regulation of LOX pathways in diatoms, as well as to perform a genetic study in support of the suggested role of lipolytic enzymes in triggering oxylipin biosynthesis in this lineage of microalgae (Figure 52).

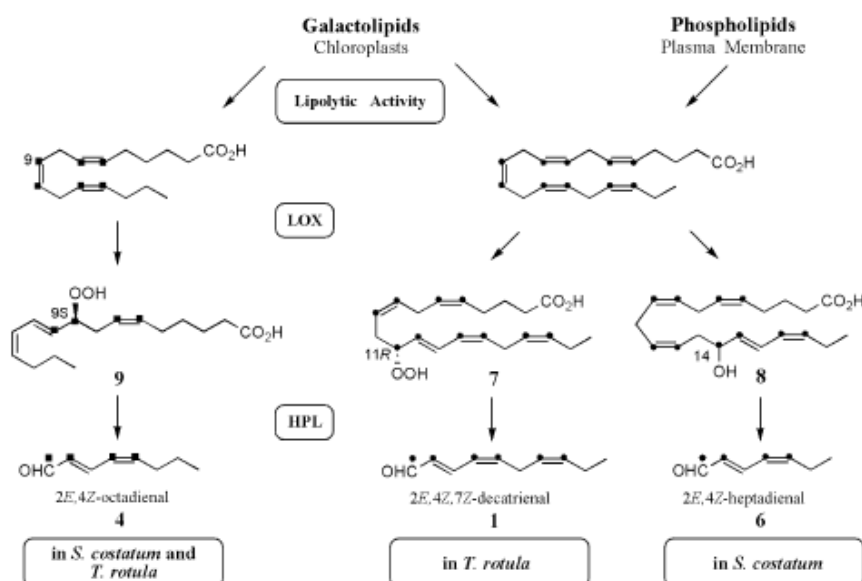


Figure 52: Model of LOX pathways in diatoms. (Fontana *et al.*, 2007. *Pure Appl. Chem.*, Vol. 79, No. 4, pp. 481–490).

During this work my first intent was the identification of genes involved in regulation of mobilization of fatty acids by hydrolysis from complex lipids under physiological conditions. An important contribution to the development of my project has been the access to the transcriptomes of

P.multistriata *P.arenysensis* and *S. marinoi*, which I choose as model species for my study.

Molecular search was performed using as query 8 genes encoding proteins belonging to the family of plant lipolytic acyl hydrolase (LAH). Among the different sequences used as queries, only the acyl hydrolase/galactolipase from *A. thaliana* (At1g61850, AtPLAI) showed a significant percentage of homology with genes of the diatoms under investigation. This protein is an acyl hydrolase involved in the mechanism of defense against pathogens and in the production of jasmonic acid. Although similar to the calcium independent mammalian phospholipase, AtPLAI is much more active on galactolipids than on phospholipids, and is localized in the chloroplast. (Yang *et al.*,2007).

In the trascriptome of *P. multistriata* my study identified a putative LAH protein of 1038 aa. The structure of the gene was also validated from access to the genome of this species recently sequenced, but not yet public. Analogously, I also isolated a putative LAH protein of 857 aa in the trascriptome of *P. arenysensis* and a sequence of 2748 bp in the trascriptome of *S. marinoi* This latter identification is still incomplete since I could not find the stop codon of the open reading frame (ORF).

From the analysis of the aligments between the sequences identified in the various transcriptomes and the sequence of *A. thaliana*, the putative

LAH proteins of *P.arenysensis* and *S.marinoi* seem to be those most similar compared to that of *P. multistriata* although all three sequences share the conserved hydrolasic motif GxSxG, and a catalytic dyad consist of Serine and Aspartic acid, elements that are critical for the LAH activity of all patatin proteins in plants (Figure 53).

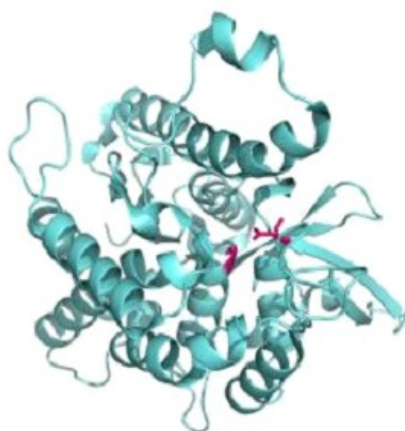


Figure 53: The 3-dimentional structure of the potato storage tuber protein patatin revealed a surface accessible Ser-Asp catalytic dyad (Ser77, Asp215, magenta) (Rydel *et al*, 2003).

In consideration that the three species of diatoms were also able to produce oxylipins by hydrolysis of fatty acids from galactolipids and phospholipids, these findings supported a functional role for these genes.

Transformation of diatoms is still an unresolved issue and so far only a few studies in *P. tricornutum* have been successfully reported (Apt *et al.*, 1996; Falciatore *et al.*, 1999). Therefore, to support the functional role of the

three putative LAHs identified in the three species of diatoms, I started a quantification of the gene expression during algal growth by qPCR techniques.

Histone H4, TBP (TATA box binding protein), RPS (30 s ribosomal protein), GAPDH (glyceraldehyde 3-phosphate dehydrogenase), tubulin (TUB A and TUB B) and ACT (actin), CDKA (cyclin dependent kinase A), COPA (coatamer A protein A) were chosen as set of reference genes in the transcriptomes of *P. multistriata*, *P. arenysensis* and *S. marinoi*

The experiments show that the expression profile of putative LAH gene normalized with H4 and RPS genes (for *P. multistriata*), COPA and CDK genes (for *P. arenysensis*), actin and ubiquitin (for *S. marinoi*) varied only slightly and randomly during the growth cycle of the three microalgae under standard conditions, thus suggesting the absence of transcriptional regulation for the three proteins.

This is very interesting since previous studies of oxylipin patterns in both *S. marinoi* (Ribalet *et al.*; 2007a, 2007b; Pohnert *et al.*, 2007) and *P. arenysensis* (d'Ippolito *et al.*, 2009) had showed modulation of synthesis of these metabolites. In fact, oxylipin levels per microalgal cell showed a progressive increase before the onset of the stationary phase in *P. multistriata* (Figure 47, Paragraph 3.5). Nevertheless the variations were too small to be significant. Similar results were also obtained with *P. arenysensis* and, in

disagreement with the literature, we could not observe a clear modulation in the synthesis of the oxylipins under the experimental conditions used during this PhD work, even if qualitative and quantitative differences (e.g levels of 15-oxoacid) were evident.

It is worthnoting that oxylipin metabolism of *S. marinoi* differs significantly from the other species for the presence of derivatives of polyunsaturated C16-fatty acids that are exclusively present in chloroplastic glycolipids (Cutignano *et al.*, 2007).

Therefore, oxylipin modulation may be dependent by many factors other than the transcriptional level of LAHs. Furthermore, the absence of significant variations in the transcriptions of the putative genes in the three species of diatoms here studied is not conclusive to exclude a regulatory function of these enzymes that, except for the fact that none of them is able to deacylate TAG, show very high variability (Huang, 1993). In fact, LAHs can work at pH between 5 and 7.5 and their activity is Ca^{2+} -independent, although the presence of this ion stimulated hydrolysis. The molecular weights of the purified proteins vary between 40 and 110 kDa and the possibility to form dimers has been also reported.

Results also suggested the presence of different LAH isozymes within the same plant material that had different substrate specificities. In fact, following that first report of Sastry *et al.* (1964) several LAHs have been

purified from different fractions of plant tissues, including soluble, chloroplastic, or microsomal fractions. Some of the lipolytic enzymes, such as those purified from the leaves of runner bean, kidney bean and potato, also had phospholipase activity. Furthermore, besides calcium stimulation, other mechanisms proposed to regulate LAH activities included phosphorylation and proteolytic activation.

The ATPLA1 (also recorded as AtPATI) sequence, that is homologous to the putative LAHs found in the three species of diatoms here studied, has a patatin domain but a much higher MW (156 kDa) than other similar proteins. AtPATI is the closest homologue to animal iPLA2s (24–32% homology) in the *Arabidopsis* genome and was found to be targeted to both the cytoplasm and chloroplasts (Holk *et al.*, 2002).

Interestingly, AtPAT1 contains a G-protein (guanine nucleotide binding protein) and the recombinant protein displays preferential galactolipase activity but also acts on phospholipids. Interestingly this enzyme also deacylates a group of oxygenated galactolipids with esterified oxylipins designated *arabidopsides* (Yang *et al.*, 2007).

The protein is also involved in synthesis of jasmonate, thus suggesting that the homologous genes we identified in diatoms may be also present in chloroplasts. In such a case, compartmentalization of the enzyme could further be a regulatory factor to take into account in the analysis of

expression of the putative LAH genes during algal growth by qPCR techniques.

All nucleus-encoded stromal plastid proteins in secondary algae appear to be preceded by an N-terminal extension, which consists of a *signal peptide* or *transit peptide* (Kroth, 2002). The characteristic feature of those presequences is the presence of a conserved motif (ASAF or AFAP). As preliminary investigation and using bioinformatic tools, the sequences of *P. multistriata*, *P. arenysensis* and *S. marinoi* were analyzed, in order to find the presence of a putative transit peptide. From the analysis of the sequences, only *S. marinoi* seems to possess some of the characteristics of transit peptide tag, including phenylalanine of the conserved motif. The analysis carried out with the software *MitoProt II* indicates the presence of a possible organelle targeting, whereas the program *TargetP 1.1* indicates the presence of a putative transit peptide that could be target proteins to a secretory pathway and a cleavage site at position 21-22. The *Predotar v. 1.03* method, instead, predicts the ER as a possible cellular localization (Figure 28; Paragraph 3.2).

This could explain why in addition to the presence of phenylalanine there is not more correspondence with the typical diatom transit peptide, probably allowing the import of the protein into the ER, but blocking the entry of the protein in the stroma. In fact this motif is generally missing at the

signal peptide cleavage site of diatom proteins putatively targeted to other destinations within the endomembrane system (e.g. into the ER).

In the sequence of *S.marinoi* it was found the motif “DAAFTS”, in which there is the conserved phenylalanine, but considering the typical conserved motif “ASAF”, there are some amino acid exchanges. In particular in the motif reported for *S.marinoi* the S has been exchanged to A (this is considered one of the exchanges allowed). After the conserved phenylalanine, the presence of the amino acids of the sequence “TS”, differs totally from the motif.

While the exchange of phenylalanine lead to blocked import, all other amino acids of the “ASAF”-motif may be replaced by glycine, alanine, serine or cysteine without affecting import. Interestingly deletions in the signal-peptide part of the motif could block plastid import.

Although large parts of the C-terminus of the transit peptide-like domain may be deleted (Apt *et al.*, 2002), plastid import is only possible if the conserved “ASAF”-motif is present between the signal and the transit peptide-like domains (Kilian & Kroth 2005). Complete deletion of either the transit peptide-like domain or the phenylalanine within the “ASAF”-motif lead to transport inhibition demonstrating that both elements are necessary. It was shown that the “ASAF”-motif and the transit peptide-like domain act as signals for actively sorting plastid proteins out of the ER/CER and for further

transport into the plastids. It has been shown that it is possible to use a signal peptide fused to “FATTP” to target the GFP (green fluorescent protein) into the plastids, while a signal peptide fused to “FA” alone fails to do the protein targeting and leads to the BLS (“blob”-like structure) phenotype (Kilian & Kroth 2005).

5. CONCLUSION AND FUTURE PERSPECTIVES

In this doctoral study, carried out at the Institute of Biomolecular Chemistry, (CNR, Pozzuoli) and at the Zoological Station “Anton Dohrn” (Naples) I worked on molecular identification of lipolytic acid hydrolases (LAHs), enzymes responsible for the release of fatty acids from phospholipids and glycolipids, in three different diatoms: *P. multistriata*, *P. arenysensis* and *S. marinoi*.

The results demonstrate for the first time the presence of putative genes homologous to the LAH gene ATPLA1 (also recorded as AtPATI) of *A. thaliana*, thus supporting previous assumptions about the involvement of galacto- and phospholipid hydrolyzing activities have in the synthesis of oxylipins in marine diatoms (d’Ippolito *et al.*, 2004; Cutignano *et al.*, 2007). Interestingly, the biochemical properties ascribed to ATPLA1 are in very well agreement with the functional evidences, such as substrate specificity, that have been circumstantially reported for the lypolitic enzymes in marine diatoms.

During the study, we also addressed transcriptional variation of these genes during growth. The results obtained from experiments did not show a clear modulation or regulation of expression of these genes under the different physiological conditions used for the cultures. Nevertheless, it has

been reported that oxylipin levels may change according to physiological or environmental factors in two of the species examined in this thesis, namely *S. marinoi* and *P. arenysensis* (Pohnert, 2008; Ribalet *et al.*, 2007; d'Ippolito *et al.*, 2009). This suggests that lipolytic enzymes do not have regulatory functions in LOX pathways of diatoms or that other mechanisms, operating after transcription, control the level of lipolytic activity. Post-transcriptional regulation of LAHs have been reported in plants, thus we cannot exclude this issue. Furthermore, the preliminary indication of a transit peptide in the putative gene coding for LAH in *S. marinoi* also suggests that compartmentalization of the protein could also play a major role in the regulation of oxylipin synthesis. To this regard, it is important to underline that diatom plastids possess a unique structure composed by four membranes that surround the organelle. Therefore, in analogy with plants where distinct LOXs are associated with different membranes in the chloroplast (Porta & Rocha-Sosa, 2002), it is possible that control of LOX pathways is simply obtained by segregation of the different enzymes in different cellular districts.

In the light of these results obtained, we will proceed to the functional characterization of the putative LAH gene in *S. marinoi*, taking into account the possible presence of the transit peptide. Currently, there is an ongoing project to express the recombinant protein in order to locate the protein

product at the subcellular level and subsequently to study its expression by antibodies.

6. REFERENCES

- Akallal R., Billard C., Fresnel J., Givernaud T. and Mouradi A. (2002). Phytoplankton of the Atlantic coasts of Morocco II. The genus *Pseudo-nitzschia* (Bacillariophyceae). *Cryptogamie Algologie*, 23: 187-202.
- Alexander H., Jenkins D.B., Ryneerson A.T., Saito A.M., Mercier L.M., Dyhrman T.S. (2012). Identifying reference genes with stable expression from high throughput sequence data. *Front. Microbiol.*, 3: 385.
- Amato A., Kooistra J.H.L., Ghiron, J.H.L., Mann D.G., Pröschold T. and Montresor M. (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist.*, 158: 193-207.
- Amato A., Orsini L., D'Alelio D., Montresor, M. (2005). Life cycle, size reduction patterns, and ultrastructure of the pennate planktonic diatom *Pseudo-nitzschia delicatissima* (Bacillariophyceae). *J. Phycol.*, 41: 542–556.
- Andersson M.X., Kourtchenko O., Dangl J.L., Mackey D., Ellerstrom M. (2006). Phospholipase-dependent signaling during the AvrRpm1- and AvrRpt2-induced disease resistance responses in *Arabidopsis thaliana*. *Plant J.*, 47: 947-959.
- Apt K.E., Kroth-Pancic P.G., Grossman A.R. (1996). Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Mol Gen Genet*, 252: 572-579

- Apt K.E., Zaslavkaia L., Lippmeier J.C., Lang M., Kilian O., Wetherbee R., Grossman A.R. and Kroth, P.G. (2002). In vivo characterization of diatom multipartite plastid targeting signals. *J. Cell Sci.*, 115: 4061–4069.

- Armbrust E.V. (2009). The life of diatoms in the world's oceans. *Nature* 459: 185-192.

- Armbrust E.V., Berges J.B., Bowler C., Green B.R., Martinez D., Putnam N.H., Zhou S., Allen A.E., Apt K.E., Bechner M., Brzezinski M.A., Chaal B.K., Chiovitti A., Davis A.K., Demarest M.S., Detter J.C., Glavina T., Goodstein D., Hadi M.Z., Hellsten U., Hildebrand M., Jenkins B.D., Jurka J., Kapitonov V.V., Kröger N., Lau V.V.Y., Lane T.W., Larimer F.W., Lippmeier J.C., Lucas S., Medina M., Montsant A., Obornik M., Schnitzler-Parker M., Palenik B., Pazour G.J., Richardson P.M., Rynearson T.A., Saito M.A., Schwartz D.C., Thamatrakoln K., Valentin K., Vardi A., Wilkerson F.P. and Rokhsar D.S. (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*, 306: 79-86.

- Armbrust E.V., Chisholm S.W. (1990). Role of light and the cell cycle on the induction of spermatogenesis in a centric diatom. *J. Phycol.*, 26: 470-478.

- Barofsky A., Vidoudez C. and Pohnert G. (2009). Metabolic profiling reveals growth stage variability in diatom exudates. *Limnol. Oceanogr. Methods*, 7: 382–390.

- Barofsky A., Pohnert G. (2007). Biosynthesis of polyunsaturated short chain aldehydes in the diatoms *Thalassiosira rotula*. *Organic Letters* 9: 1017–1020.

- Bhattacharya D., Medlin L. (1995). The phylogeny of plastids: a review based on comparison of small subunit ribosomal coding region. *J. Phycol.*, 31: 489-498.

- Bhaya D., Grossman A. (1991). Targeting proteins to diatom plastids involves transport through an endoplasmic reticulum. *Mol. Gen. Genet.*, 229: 400–404.

- Bender S.J., Parker M.S. and Armbrust E.V. (2012). Coupled effects of light and nitrogen source on the urea cycle and nitrogen metabolism over a diel cycle in the marine diatom *Thalassiosira pseudonana*. *Protist*, 163: 232–251.

- Berg G.M., Shrager J., Glöckner G., Arrigo K. R., and Grossman A. R. (2008). Understanding nitrogen limitation in *Aureococcus anophagefferens* (*Pelagophyceae*) through cDNA and qRT-PCR analysis. *J. Phycology*, 44 (5): 1235-1249.

- Blée E. (2002). Impact of phyto-oxylipins in plant defence. *Trends in Plant Science*. 7: 315-321.

- Bouck, G.B. (1965). Fine structure and organelle associations in brown algae. *J. Cell Biol.*, 26: 523-537.

- Bowler C. (2008). Diatom signalling: deadly messages. *Curr. Biol.*, 18 (12): 518-519.

- Bowler C., Allen A.E., Badger J.H., Grimwood J., Jabbari K., Kuo A., Maheswari U., Martens C., Maumus F., Otillar R.P., Rayko E., Salamov A., Vandepoele K., Beszteri B., Gruber A., Heijde M., Katinka M., Mock T., Valentin K., Verret F., Berges J.A., Brownlee C., Cadoret J.P., Chiovitti A., Choi C.J., Coesel S., De Martino A., Detter J.C., Durkin C., Falciatore A., Fournet J., Haruta M., Huysman M.J., Jenkins B.D., Jiroutova K., Jorgensen R.E., Joubert Y., Kaplan A., Kröger N., Kroth P.G., La Roche J., Lindquist E., Lommer M., Martin-Jézéquel V., Lopez P.J., Lucas S., Mangogna M., McGinnis K., Medlin L.K., Montsant A., Oudot-Le Secq M.P., Napoli C., Obornik M., Parker M.S., Petit J.L., Porcel B.M., Poulsen N., Robison M., Rychlewski L., Ryneerson T.A., Schmutz J., Shapiro H., Siaut M., Stanley M., Sussman M.R., Taylor A.R., Vardi A., von Dassow P., Vyverman W., Willis A., Wyrwicz L.S., Rokhsar D.S., Weissenbach J., Armbrust E.V., Green B.R., Van de Peer Y., Grigoriev I.V. (2008). The *Phaeodactylum* genome reveals the evolutionary history of diatom genome. *Nature*, 456 (7219): 239-44.

- Bruce B.D. (2000). Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol.*, 10: 440–447.

- Bustin S.A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. *J. Mol. Endocrinol.* 29: 23–39.

- Bustin S.A., Beaulieu J.F., Huggett J., Jaggi R., Kibenge, F.S., Olsvik, P.A., Penning L.C., Toegel S. (2010). MIQE precis: practical implementation of minimum standard guidelines for fluorescencebased quantitative real-time PCR experiments. *BMC Molecular Biology*, 11: 74.

- Cannell, R.J.P. (1998). How to approach the isolation of a natural product, p. 1–51. In R.P.J. Cannell (ed.), *Natural Products Isolation, Methods in Biotechnology*, Volume 4. Humana Press, Totowa, New Jersey.

- Cao S., Zhang X., Ye N., Fan X., Mou S., Xu D., Liang C., Wang Y., Wang W. (2012). Evaluation of putative internal reference genes for gene expression normalization in *Nannochloropsis sp.* by quantitative real-time RT-PCR. *Biochem. Biophys. Res. Commun.*, 424 (1): 118-123.

- Casotti R., Mazza S., Brunet C., Vantrepotte V., Ianora A., Miralto A. (2005). Growth inhibition and toxicity of the diatom aldehyde 2-trans-4-trans decadienal on *Thalassiosira weissflogii* (Bacillariophyceae). *Journal of Phycology*, 41: 7–20.

- Cavalier-Smith T. (2000). Membrane heredity and early chloroplast evolution. *Trends Plant Sci.*, 5: 174–182.

- Cavalier-Smith T. (2003). Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote- eukaryote chimaeras (meta-algae). *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 358: 109–134.

- Chandra S., Heinsteins P.F., Low P.S. (1996). Activation of phospholipase a by plant defense elicitors. *Plant Physiol.*, 110: 979–986.

- Chapman K.D. (1998). Phospholipase activity during plant growth and development and in response to environmental stress. *Trends Plant Sci.*, 3: 419–426.

- Chepurinov V.A., Mann D.G., Sabbe K., Vannerum K., Casteleyn G., Verleyen E., Peperzak L., Vyverman W. (2005). Sexual reproduction, mating system, chloroplast dynamics and abrupt cell size reduction in *Pseudo-nitzschia pungens* from the North Sea (*Bacillariophyta*). *Eur. J. Phycol.*, 40: 379–395.

- Cho E.S., Hur H.J., Byun H.S., Lee S.G., Rhodes L.L., Jeong C.S., Park J.G. (2002). Monthly monitoring of domoic acid producer *Pseudo-nitzschia multiseries* (Hasle) using species-specific DNA probes and WGA lectins and abundance of *Pseudo-nitzschia species* (*Bacillariophyceae*) from Chinae Bay. *Korea Bot. Mar.*, 45: 364–372.

- Colin S.P., Dam H.G. (2002). Testing for toxic effects of prey on zooplankton using sole versus mixed diets. *Limnol. Oceanogr.*, 47: 1430–1437.

- Cronin G. (2001). Resource allocation in seaweeds and marine invertebrates: Chemical defense patterns in relation to defense theories, p. 325–353. In J. B. McClintock and B. J. Baker (eds.), *Marine Chemical Ecology*. CRC Press, Boca Raton, Florida.

- Cutignano A., d'Ippolito G., Romano G., Cimino G., Febbraio F., Nucci R., Fontana A. (2006). Chloroplastic galactolipids fuel the aldehyde biosynthesis in the marine diatom *Thalassiosira rotula*. *ChemBioChem.*, 7 (3): 450-456.

- Cutignano A., Lamari N., d'Ippolito G., Manzo E., Cimino G., Fontana A. (2011). Lipoxygenase products in marine diatoms: a concise analytical method to explore the functional potential of oxylipins. *Journal of Phycology* Volume 47, Issue 2, pages 233–243.

- Davidovich N.A., Bates S.S. (1998). Sexual reproduction in the pennate diatoms *Pseudo-nitzschia multiseriata* and *P. pseudodelicatissima* (Bacillariophyceae). *J. Phycol.*, 34: 126–137.

- Davis A.K., Hildebrand M. and Palenik B. (2006). Gene expression induced by copper stress in diatom *Thalassiosira pseudonana*. *Eukaryot. Cell*, 5: 1157–1168. - Davis A.K. and Palenik B. (2008). Characterization of a modular, cell surface, protein and identification of a new gene family in the diatom *Thalassiosira pseudonana*. *Protist*, 159: 195–207.

- Dhondt S., Geoffroy P., Stelmach B.A., Legrand M. and Heitz, T. (2000). Soluble phospholipase A2 activity is induced before oxylipin accumulation in tobacco mosaic virus-infected tobacco leaves and is contributed by patatin-like enzymes. *Plant J.*, 23: 431–440.

- Dhondt S., Gouzerh G., Muller A., Legrand M. and Heitz T. (2002). Spatio-temporal expression of patatin-like lipid acyl hydrolases and accumulation of jasmonates in elicitor-treated tobacco leaves are not affected by endogenous levels of salicylic acid. *Plant J.*, 32: 749–762.

- d'Ippolito G., Cutignano A., Tucci S., Romano G., Cimino G., Fontana A. (2006). Biosynthetic intermediates and stereochemical aspects of aldehyde biosynthesis in the marine diatom *Thalassiosira rotula*. *Phytochemistry*, 67: 314–322.

- d'Ippolito G, Cutignano A, Romano G, Lamari N, Massa Gallucci A, Cimino G, Miralto A, Ianora A., Fontana A. (2007). LOX-induced lipid peroxidation mechanism responsible for the detrimental effect of marine diatoms on zooplankton grazers. *Chembiochem*. 2007 ;8(15):1810-8.

- d'Ippolito G., Iadicicco I., Romano G., Fontana A. (2002b). Detection of short-chain aldehydes in marine organisms: the diatom *Thalassiosira rotula*. *Tetrahedron Lett.*,43: 6137–6140.

- d'Ippolito G., Romano G., Caruso T., Spinella A., Cimino G., Fontana A. (2003). Production of octadienal in the marine diatom *Skeletonema costatum*. *Org. Lett.*, 5: 885–887.

- d'Ippolito G., Romano G., Iadicicco O., Miralto A., Ianora A., Cimino G., Fontana A. (2002a) New birth-control aldehydes from the marine diatom *Skeletonema costatum*: characterization and biogenesis. *Tetrahedron Lett.*, 43: 6133–6136.

- Domergue F., Spiekermann P., Lerchl J., Beckmann C., Kilian O., Kroth P.G., Boland W., Zahringer U. and Heinz E. (2003). New insight into *Phaeodactylum tricornutum* fatty acid metabolism. Cloning and functional characterization of plastidial and microsomal delta12-fatty acid desaturases. *Plant Physiol.*, 131: 1648–1660.

- Dyhrman S.T. (2011). Proteome changes driven by phosphorus deficiency and recovery in the brown tide-forming alga *Aureococcus anophagefferens*. *PLoS ONE*, 6:e28949. doi: 10.1371/journal.pone.0028949.

- Enkhbayar P., Kamiya M., Osaki M., Matsumoto T. and Matsushima N. (2004). Structural principles of leucine-rich repeat (LRR). *Proteins*, 54: 394-403.

- Falciatore A., Casotti R., Leblanc C., Abrescia C., Bowler C. (1999). Transformation of nonselectable reporter genes in marine diatoms. *Mar Biotechnol*, 1: 239-251.

- Feeny P. (1992). The evolution of chemical ecology: Contributions from the study of herbivorous insects, p. 1–44. In G. A. Rosenthal and M. R. Berenbaum (eds.), *Herbivores: Their Interactions with Secondary Plant Metabolites, Volume II: Ecological and Evolutionary Processes*. Academic Press, San Diego, California.

- Ferrante M.I., Romio L., Castro S., Collins J.E., Goulding D.A., Stemple D.L., Woolf A.S., Wilson S.W. (2009). Convergent extension movements and ciliary function are mediated by ofd1, a zebrafish orthologue of the human oral-facial-digital type 1 syndrome gene. *Hum Mol Genet*. 15;18(2):289-303. doi: 10.1093/hmg/ddn356

- Feussner I., Kuhn H., Wasternack C. (2001) Lipoxygenase-dependent degradation of storage lipids. *Trends Plant Sci.*, 6: 268–273.

- Field C.B., Behrenfeld M.J., Randerson J.T. and Falkowski P. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science - New York then Washington* Vol. 281.

- Fischer A.M., Dubbs W.E., Baker R.A., Fuller M.A., Stephenson L.C., Grimes H.D. (1999). Protein dynamics, activity and cellular localization of soybean lipoxygenases indicate distinct functional roles for individual isoforms. *The Plant Journal*, 19 (5): 543-554.

- Fontana A., d'Ippolito G., Cutignano A., Miralto A., Ianora A., Romano G., Cimino G. (2007a) Chemistry of oxylipin pathways in marine diatoms. *Pure Appl. Chem.*, 79: 475–484.

- Fontana A., d'Ippolito G., Cutignano A., Romano G., Lamari N., Massa Gallucci A., Cimino G., Miralto A., Ianora A (2007b). A metabolic mechanism for the detrimental effect of marine diatoms on zooplankton grazers. *ChemBioChem.*, 8: 1810–1818.

- Frixell, G.A., Hasle G.R. (2003). Taxonomy of harmful diatoms, p. 465–509. In G.M. Hallegraeff, D.M. Anderson, and A.D. Cembella [eds.], *Manual on harmful marine microalgae. UNESCO*.

- Gibbs S.P. (1979). The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *J. Cell Sci.*, 35: 253–266.

- Goss R., Nerlich J., Lepetit B., Schaller S., Vieler A., Wilhelm C. (2009). The lipid dependence of diadinoxanthin de-epoxidation presents new evidence for macrodomain organization of the diatom thylakoid membrane. *Journal of Plant Physiology*, 166: 1839–1854.

- Goss R., Wilhelm C. (2009). Lipids in algae, lichens and mosses. In Wada J., Murata N., eds. *Advances in photosynthesis, lipids in photosynthesis: essential and regulatory functions* 29. *Springer Science and Business Media*, 117–137.

- Grienemberger E., Geoffroy P., Mutterer J., Legrand M., Heitz T. (2010). The interplay of lipid acyl hydrolases in inducible plant defense. *Plant Signaling & Behavior* 5, 10: 1181-1186.

- Hamm C.E., Merkel R., Springer O., Jurkojc P., Maier C., Prechtel K., Smetacek V. (2003). Architecture and material properties of diatom shells provide effective mechanical protection. *Nature*, 421: 841–843.

- Harper M.K., Bugni T.S., Copp B.R., James R.D., Linsay B.S., Richardson A.D., Schnabel P.C., Tesdemir D., Vanwagoner R.M., Verbitiski S.M., Ireland C.M. (2001). Introduction to the chemical ecology of marine natural products, p. 3–29. In J. B. McClintock and B. J. Baker (eds.).

- Hasle G.R. (2002). Are most of the domoic acid-producing species of the diatom genus *Pseudo-nitzschia* cosmopolites? *Harmful Algae*, 1: 137-146.

- Hasle G.R., Syvertsen E.E. (1997). Marine diatoms. In: Tomas, C.R. (Ed.), *Identifying Marine Phytoplankton*. *Academic Press*, San Diego, pp. 5–385.

- Hempel F., Felsner G., Maier G.U. (2010). New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms *Molecular Microbiology*, 76 (3): 793–801.

- Hendriks T., Vreugdenhil D., Stiekema W.J.. (1991). Patatin and four serine proteinase inhibitor genes are differentially expressed during potato tuber development. *Plant Mol. Biol.*, 17: 385-94.

- Herbert R.B. (1989). *The Biosynthesis of Secondary Metabolites* (Chapman & Hall).

- Holk A., Rietz S., Zahn M., Quader H., Scherer G.F. (2002). Molecular identification of cytosolic, patatin-related phospholipases A from *Arabidopsis* with potential functions in plant signal transduction. *Plant Physiol.*, 130: 90-101.

- Hombeck M., Boland W. (1998). Biosynthesis of the algal pheromone fucoserratene by the freshwater diatom *Asterionella Formosa* (Bacillariophyceae). *Tetrahedron* 54: 11033–11042.

- Hombeck M, Pohnert G, Boland W. (1999). Biosynthesis of dictyopterene A: stereoselectivity of a lipoxygenase/hydroperoxide lyase from *Gomphonema parvulum* (Bacillariophyceae). *Journal of the Chemical Society Chemical Communications* 3: 243–244.

- Howe G.A., Schilmiller A.L. (2002). Oxylin metabolism in response to stress. *Curr. Opin.Plant Biol.*, 5 (3): 230-236.

- Huang A.H.C., Lipases, in: T.S. Moore Jr (1993). (Ed.), *Lipid Metabolism in Plants*, CRC Press, Boca Raton, FL , pp. 473–503.

- Hughes R.K., Yousafzai F.K., Ashton R., Chechetkin I.R., Fairhurst S.A., Hamberg M., Casey R. (2008). Evidence for communality in the primary determinants of CYP74 catalysis and of structural similarities between CYP74 and classical mammalian P450 enzymes. *Proteins* 72: 1199–1211

- Ianora A., Boersma M., Casotti R., Fontana A., Harder J., Hoffmann F., Pavia H., Potin P., S. Poulet A., Toth G. (2006) New Trends in Marine Chemical Ecology. *Estuaries and Coasts*, 29 (4): 531–551.

- Ianora A., Miralto A., Poulet S.A., Carotenuto Y., Buttino I., Romano G., Casotti R., Pohnert G., Wichard T., Colucci-D’Amato L., Terrazzano G., Smetacek V. (2004). Aldehyde suppression of copepod recruitment in bloom of ubiquitous planktonic diatom. *Nature*, 429: 403-407.

- Inderjit, Duke S.O. (2003). Ecophysiological aspects of allelopathy. *Planta* 217: 529–539.

- Ishida K., Cavalier-Smith T., Green B.R. (2000) Endomembrane structure and the chloroplast protein targeting pathway in *Heterosigma akashiwo* (*Raphidophyceae*, *Chromista*). *J. Phycol.*, 36: 1135–1144.

- Jackson-Constan D., Akita M. and Keegstra K. (2011). Molecular chaperones involved in chloroplast protein import. *Biochim. Biophys. Acta*, 1541: 102–113.

- Kaczmarska I., Bates S.S., Ehrman J.M., Leger C. (2000). Fine structure of the gamete, auxospore and initial cell in the pennate diatom *Pseudo-nitzschia multiseries* (Bacillariophyta). *Nova Hedwigia*, 71: 337–357.

- Kaczmarska I., LeGresley M.M., Martin J.L., Ehrman J. (2005). Diversity of the diatom genus *Pseudo-nitzschia* Peragallo in the Quoddy Region of the Bay of Fundy, Canada. *Harmful Algae*, 4: 1–19.

- Keeling P.J., Burger G., Durnford G.D., Lang B.F., W. Lee R., Pearlman E.R., Roger J.A., Gray W.M. (2005) *TRENDS in Ecology and Evolution*, 20: 670.

- Kemp A.E.S., Pearce R.B., Grigorov I., Rance J., Lange C.B., Quilty P. and I.S. (2006). Production of giant marine diatoms and their export at oceanic frontal zones: implications for Si and C flux from stratified oceans. *Global Biogeochem Cycles* 20.

- Kienesberger P.C., Oberer M., Lass A., Zechner R. (2009). Mammalian patatin domain containing proteins: a family with diverse lipolytic activities involved in multiple biological functions. *J. Lipid Res.*, 50: 63-68.

- Kilian O., Kroth P.G. (2005). Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *The Plant Journal*, 41: 175–183.

- Kilian O., Kroth P.G. (2003) Evolution of protein targeting into ‘complex’ plastids: the ‘secretory transport hypothesis’. *Plant Biol.*, 5: 350–358.

- Kilian O., Kroth P.G. (2004). Presequence acquisition during secondary endocytobiosis and the possible role of introns. *J. Mol. Evol.*, 58: 712–721.

- Kolomiets M.V., Hannapel D.J., Chen H., Tymeson M., Gladon R.J. (2001) Lipoxygenase is involved in the control of potato tuber development. *The Plant Cell*, 13: 613–626.

- Kroth P.G. (2002) Protein transport into secondary plastids and the evolution of primary and secondary plastids. *Int. Rev. Cytol.*, 221: 191–255.

- Kunau W.H. (2001). Peroxisomes: the extended shuttle to the peroxisome matrix. *Curr. Biol.*, 11: R659–R662.
- Kuo K.L., Feng-Hsiu Tsua H.F., Chang J. (2012). Quantification of diatom gene expression in the sea by selecting uniformly transcribed mRNA as the basis for normalization. *Appl. Environ. Microbiol.*, doi:10.1128/AEM.00935.

- Kwon S.J., Jin H.C., Lee S., Nam M.H., Chung J.H., Kwon S.I., Ryu C.M., Park O.K. (2009). GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in *Arabidopsis*. *Plant J.*, 58: 235-245.

- La Camera S., Geoffroy P., Samaha H., Ndiaye A., Rahim G., Legrand M., Heitz T. (2005). A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in *Arabidopsis*. *Plant J.*, 44: 810-825.

- La Camera S., Gouzerh G., Dhondt S., Hoffmann L., Fritig B., Legrand M., Heitz T. (2004). Metabolic reprogramming in plant innate immunity: the contributions of phenylpropanoid and oxylipin pathways. *Immunol. Rev.*, 198: 267-284.

- Landsberg J.H. (2002). The effects of harmful algal blooms on aquatic organisms. *Rev. Fish. Sci.*, 10: 113–390.

- Lang M., Apt K.E., Kroth P.G. (1998). Protein transport into ‘complex’ diatom plastids utilizes two different targeting signals. *J. Biol. Chem.*, 273: 30973–30978.

- Le Bail A., Dittami M.S., de Franco O.P., Rousvoal S., Cock J.M., Tonon T., Charrier B. (2008). Normalization genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. *BMC Molecular Biology*, 9: 75.

- Leblanc C., Falciatore A., Bowler C. (1999). Semi-quantitative RT-PCR analysis of photoregulated gene expression in marine diatoms. *Plant Mol. Biol.*, 40: 1031-1044.

- Lee D.S., Nioche P., Hamberg M., Raman C.S. (2008). Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. *Nature* 455: 363–368.

- Legendre L. (1990). The significance of microalgal blooms for fisheries and for the export of particulate organic carbon in oceans. *J. Plankton Res.*, 12: 681–699.

- Lepetit B., Goss R., Jakob T., Wilhelm C. (2011). Molecular dynamics of the diatom thylakoid membrane under different light conditions. *Photosynthesis Research*, doi: 10.1007/s11120-011-9633-5.

- Liu C., Wu G., Huang X., Liu S., Cong B. (2012). Validation of housekeeping genes for gene expression studies in an ice alga *Chlamydomonas* during freezing acclimation. *Extremophiles*, 16 (3): 419-25.

- Long J. D., Smalley G. W., *et al.* (2007). "Chemical cues induce consumer-specific defenses in a bloom-forming marine phytoplankton." *Proceedings of the National Academy of Sciences of the United States of America* 104(25): 10512-10517.

- Maneiro I., Iglesias P., Guisande C., Riveiro I., Barreiro A., Zeroudaki S., Granéli E. (2005). Fate of domoic acid ingested by the copepod *Acartia clausi*. *Mar. Biol.*, 148: 123–130.

- Mann D.G. (1993). Patterns of sexual reproduction in diatoms. *Hydrobiologia*, 269/270: 11-20.

- Mannitto P.(1981), *Biosynthesis of Natural Products Marine Chemical Ecology*. CRC Press, Boca Raton, Florida.

- Martin W., Rujan, T., Richly E., Hansen A., Cornelsen S., Lins T., Leister D., Stoebe B., Hasegawa M., Penny D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. USA* 99, 429–441.

- Matos A.R., Pham-Thi A.T. (2009). Lipid deacylating enzymes in plants: old activities, new genes. *Plant Physiol. Biochem.*, 47: 491-503.

- Matsui K., Kurishita S., Hisamitsu A., Kajiwara T. (2000). A lipid-hydrolyzing activity involved in hexenal formation. *Biochem. Soc. Trans.*, 28: 857–860.

- McClintock J.B., Baker B.J. (2001). *Marine Chemical Ecology Series*.

- McFadden G.I. (1999). Plastids and protein targeting. *J. Eukaryot. Microbiol.*, 46: 339–346.

- McGinn P.J. and Morel F.M.M. (2008a). Expression and inhibition of the carboxylating and decarboxylating enzymes in photosynthetic C4 pathway of marine diatoms. *Plant Physiol.*, 146: 300–309.

- McGinn P.J. and Morel F.M.M. (2008b). Expression and regulation of carbonic anhydrases in the marine diatom *Thalassiosira pseudonana* and in natural phytoplankton assemblages from Great Bay, New Jersey. *Physiol. Plant*, 133: 78–91.

- Miralto A., Barone G., Romano G., Poulet S.A., Ianora A., Russo G.L., Buttino I., Mazzarella G., Laabir M., Cabrini M., Giacobbe M.G. (1999). The insidious effect of diatoms on copepod reproduction. *Nature*, 402: 173–176.

- Miralto A., Ianora A., Guglielmo L., Zagami G., Buttino I. (1998). Egg production and hatching success in the peri-Antarctic copepod *Calanus simillimus*. *J. Plankton Res.*, 20: 2369–2378.

- Mock T., Samanta M.P., Iverson V., Berthiaume C., Robison M., Holtermann K., Durkin C., Bondurant S.S., Richmond K., Rodesch M., Kallas T., Huttlin E.L., Cerrina F., Sussman M.R., Ambrust E.V. (2008). Whole-genome expression pro-filing of the marine diatom *Thalassiosira pseudonana* identifies genes involved in silicon bioprocesses. *Proc. Natl. Acad. Sci. U.S.A.*, 105: 1579–1584.

- Moestrup Ø (ed) (2004) IOC Taxonomic Reference List of Toxic Plankton Algae, Intergovernmental Oceanographic Commission of UNESCO; www.bi.ku.dk/ioc/default.asp

- Moseley J.L., Chang C.W. and Grossman A.R. (2006). Genome-based approaches to understanding phosphorus deprivation responses and PSR1 control in *Chlamydomonas reinhardtii*. *Eukaryot. Cell*, 5: 26–44.

- Nassoury N., Cappadocia M., Morse D. (2003). Plastid ultrastructure defines the protein import pathway in dinoflagellates. *J. Cell Sci.*, 116: 2867–2874.

- Nejstgaard J.C., Solberg P.T. (1996). Repression of copepod feeding and fecundity by the toxic haptophyte *Prymnesium patelliferum*. *Sarsia*, 81: 339–344.

- Neuhaus J. and Rogers J.C. (1998). Sorting of proteins to vacuoles in plant cells. *Plant Mol. Biol.*, 38: 127–144.

- Norton A.T., Melkonian M., Andersen A.R. (1996) Algal biodiversity. *Phycologia*, 35 (4): 308-326.

- Oh I.S., Park A.R., Bae M.S., Kwon S.J., Kim Y.S., Lee J.E., Kang N.Y., Lee S., Cheong H., Park O.K. (2005). Secretome analysis reveals an *Arabidopsis* lipase involved in defense against *Alternaria brassicicola*. *Plant Cell*, 17: 2832-2847.

- Orive E., Laza-Martinez A., Seoane S., Alonso A., Andrade R. and Miguel I. (2010). Diversity of *Pseudo-nitzschia* in the southeastern Bay of Biscay. *Diatom Research*, 25: 125-145.

- Orsini L., Sarno D., Procaccini G., Poletti R., Dahlmann J., Montresor M. (2002). Toxic *Pseudo-nitzschia multistriata* (Bacillariophyceae) from the Gulf of Naples: morphology, toxin analysis and phylogenetic relationships with other *Pseudonitzschia* species. *Eur. J. Phycol.*, 37: 247–257.

- Pfaffl W.M. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleid Acid Research*, Volume 29, Issue 9, No 9.

- Pfanner N. and Geissler A. (2001). Versatility of the mitochondrial protein import machinery. *Nature Rev. Mol. Cell. Biol.*, 2: 339–349.

- Pohnert G. (2000). Wound-activated chemical defense in unicellular planktonic algae. *Angew. Chem. Int.*, 39: 4352–4354.

- Pohnert G. (2002). Phospholipase A₂ activity triggers the wound-activated chemical defense in the diatom *Thalassiosira rotula*. *Plant Physiology*, 129: 103–111.

- Pohnert G., Boland W. (2002). The oxylipin chemistry of attraction and defense in brown algae and diatoms. *Natural Product Reports* 19: 108–122

- Pohnert G., Steinke M., Tollrian R. (2007) Chemical cues, defence metabolites and the shaping of pelagic interspecific interactions. *Trends Ecol. Evol.*, 22, 198–204.

- Porta H., Rocha-Sosa M. (2002). Plant Lipoxygenases. Physiological and Molecular Features. *Plant Physiology*, 130: 15–21.

- Raffaele S., Leger A., Roby D. (2009). Very long chain fatty acid and lipid signaling in the response of plants to pathogens. *Plant Signal Behav.*, 4: 94-99.

- Ribalet F., Berges J. A., Ianora A. *et al.* (2007a) Growth inhibition of cultured marine phytoplankton by toxic algal-derived polyunsaturated aldehydes. *Aquatic Toxicol*, 85, 219–227.

- Ribalet F., Wichard T., Pohnert G. *et al.* (2007b) Age and nutrient limitation enhance polyunsaturated aldehyde production in marine diatoms. *Phytochemistry*, 68, 2059–2067.

- Roberts M.F. (1981). Enzymic synthesis of γ -coniceine in *Conium maculatum*. *Plant Cell Rep.*, 1: 10-13.

- Rosic N.N., Pernice M., Dove S., Dunn S. and Hoegh-Guldberg O. (2010a). Gene expression profiles of cytosolic heat shock proteins Hsp70 and Hsp90 from symbiotic dinoflagellates in response to thermal stress: possible implications for coral bleaching. *Cell Stress Chaperon.*, 16: 69–80.

- Rosic N.N., Pernice M., Rodriguez Lanetty M. and Hoegh-Guldberg, O. (2010b). Validation of housekeeping genes for gene expression studies in *Symbiodinium* exposed to thermal and light stress. *Mar. Biotechnol.*, 13: 355–365.

- Round F.E., Crawford R.M. and Mann D.G. (1990). *The Diatoms. Biology and Morphology of the Genera*, Cambridge University Press.

- Ruoyu G., Jang S.K. (2012). Evaluation and validation of internal control genes for studying gene expression in the dinoflagellate *Prorocentrum minimum* using real-time PCR. *European Journal of Protistology*, 48: 199–206.

- Ruoyu G., Min-Ah Lee, Jang-Seu Ki. (2012). Normalization genes for mRNA expression in the marine diatom *Ditylum brightwellii* following exposure to thermal and toxic chemical stresses. *Journal of Applied Phycology*, doi 10.1007/s10811-012-9908-z.

- Rydel T.J., Williams J.M., Krieger E., Moshiri F., Stallings W.C., Brown S.M., Pershing J.C., Purcell J.P., Alibhai, M.F. (2003). The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad. *Biochemistry*, 42: 6696–6708.

- Ryu S.B. (2004). Phospholipid-derived signaling mediated by phospholipase A in plants. *Trends Plant Sci.*, 9: 229–235.
- Sarno D. and Dahlman J. (2000). Production of domoic acid in another species of *Pseudo-nitzschia*: *P. multistriata* in the Gulf of Naples (Mediterranean Sea). *Harmful Algae News*, 21: 5.

- Sastry P.S., Kates M. (1964). Hydrolysis of monogalactosyl diglycerides by specific enzymes in runner bean leaves, *Biochemistry* 3 1280–1287.

- Scherer G.F., Ryu S.B., Wang X., Matos A.R., Heitz T. (2010). Patatin-related phospholipase A, subfamilies and functions in plants and animals. *Trends Plant Sci.*, 15: 693-700.

- Schwarz K., Walther M., Anton M., Gerth C., Feussner I., Kuhn H. (2001). Structural basis for lipoxygenase specificity. Conversion of the human leukocyte 5-lipoxygenase to a 15-lipoxygenating enzyme species by site-directed mutagenesis. *J. Biol. Chem.*, 5;276(1):773-9.

- Sellner K.G., Olson M.M., Olli K. (1996) Copepod interactions with toxic and non-toxic cyanobacteria from the Gulf of Finland. *Phycologia*, 35: 177–182.

- Shah J. (2005). Lipids, lipases and lipid-modifying enzymes in plant disease resistance. *Annu Rev Phytopathol.*, 43: 229-260.

- Shaw B.A., Andersen R.J., Harrison P.J. (1997). Feeding deterrent and toxicity effects of apo-fucoxanthinoids and phycotoxins on a marine copepod (*Tigriopus californicus*) *Mar Biol.*, 128: 273–280.

- Siaut M., Heijde M., Mangonga M., Montsant A., Coesel S., Allen A., Falciatore A., Bowler C. (2007). Molecular toolbox for studying diatom biology in *Phaeodactylum tricornutum*. *Gene*, 406 (1-2): 23-35.

- Siedow J.N. (1991). Plant lipoxygenase: structure and function. *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, 42: 145–188.

- Sloane D.L., Leung R., Craik C.S., Sigal E. (1991). A primary determinant for lipoxygenase positional specificity. *Nature*, 14;354(6349):149-52

- Sloane D.L., Leung R., Barnett J., Craik C.S., Sigal E. (1995). Conversion of human 15-lipoxygenase to an efficient 12-lipoxygenase: the side-chain geometry of amino acids 417 and 418 determine positional specificity. *Protein Eng.*, 8(3):275-82

- Smetacek V. (1999). Diatoms and the ocean carbon cycle. *Protist*, 150: 25-32.

- Soll S., Schleiff E. (2004). Proteins import into chloroplast. *Nature Reviews Molecular Cell Biology*, 5: 198–208.

- Steinert M., Hentschel U., Hacker J. (2000). Symbiosis and pathogenesis: Evolution of the microbe-host interaction. *Naturwissenschaften*, 87: 1–11.

- Steinke M., Malin G., Liss P.E. (2002). Trophic interactions in the sea: An ecological role for climate relevant volatiles. *Journal of Phycology*, 38: 630–638.

- Stelmach B.A., Muller A., Hennig P., Gebhardt S., Schubert-Zsilavecz M., Weiler E.W. (2001). A novel class of oxylipins, sn1-O-(12-oxophytodienoyl)-sn2-O-(hexadecatrienoyl)-monogalactosyl diglyceride, from *Arabidopsis thaliana*. *J. Biol. Chem.*, 276: 12832–12838.

- Stuart R.K., Dupont C.L., Johnson D.A., Paulsen I.T. and Palenik B. (2009). Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl. Environ. Microbiol.*, 75: 5047–5057.

- Svesnikova N., Soll J. and Schleiff E. (2000). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc. Natl Acad. Sci. USA*, 97: 4973–4978.

- Takano H. (1993). Marine diatom *Nitzschia multistriata* sp. nov. common at inlets of southern Japan. *Diatom*, 8: 39–41.

- Takano H. (1995). *Pseudo-nitzschia multistriata* (Takano) Takano, a new combination for the pennate diatom *Nitzschia multistriata* Takano. *Diatom*, 10: 73–74.

- Tesson S.V.M. (2010). Population genetic structure of a planktonic diatom in the Gulf of Naples: *Pseudo-nitzschia multistriata*. PhD thesis

- Turner J.T., Tester P.A., Hansen P.J. (1998). Interactions between toxic marine phytoplankton and metazoan and protistan grazers. In *Physiological Ecology of Harmful Algal Blooms*. Anderson, D.M., Cembella, A.D., Hallegraeff, G.M., Eds., Springer-Verlag: Berlin, Germany: 453–474.

- Van Den Hoek C., Mann D.G., Jahns H.M. (1997). *Algae. An Introduction to Phycology*. Cambridge University Press.

- Vardi A., Bidle K., Kwityn C., Hirsh D.J., Thompson S.M., Callow. J.A., Falkowski P., Bowler C. (2008). A diatom gene regulating Nitric-Oxide signalling and susceptibility to diatom-derived aldehydes. *Current Biology*, 18 (12): 895-9.

- Vidoudez C., Pohnert G. (2008). Growth phase-specific release of polyunsaturated aldehydes by the diatom *Skeletonema marinoi*. *J. Plankton Res.*, 30: 1305-1313.

- Vieler A., Sub R., Wilhelm C., Schiller J. (2007). The lipid composition of the unicellular green alga *Chlamydomonas reinhardtii* and the diatom *Cyclotella meneghiniana* investigated by MALDI-TOF MS and TLC. *Chemistry and Physics of Lipids*, 150: 143–155.

- Villareal, T.A., Carpenter, E.J. (1989). Nitrogen fixation, suspension characteristics and chemical composition of *Rhizosolenia mats* in the central North Pacific gyre. *Biol. Oceanogr.*, 6: 327-345.

- Wang X.M. (2001). Plant phospholipases. *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, 52: 211–231.

- Wendel T., Jüttner F. (1996). Lipoxygenase-mediated formation of hydrocarbons and unsaturated aldehydes in freshwater diatoms. *Phytochemistry*, 6: 1445–1449.

- Whitney L.P., Lins J.J., Chappell P.D. and Jenkins B.D. (2011). Characterization of putative iron responsive genes as species-specific indicators of iron stress in *Thalassiosira* diatoms. *Front. Microbio.*, 2: 234.

- Wichard T., Pohnert G. (2006). Formation of halogenated medium chain hydrocarbons by a lipxygenase/hydroperoxide halolyase-mediated transformation in planktonic microalgae. *Journal of the American Chemical Society* 128: 7114–7115

- Williams D.H., Stone M.J., Hauck P.R., Rahman S.K. (1989). Why are secondary metabolites (natural products) biosynthesized? *Journal of Natural Products*, 52: 1189–1208.

- Wink M., Hartmann T., Schiebel H.M. (1979). A model mechanism for the enzymatic synthesis of lupin alkaloids. *Z. Naturforsch*, 35c: 93-97.

- Wolfe G.V. (2000). The chemical defense ecology of marine unicellular plankton: Constraints, mechanisms, and impacts. *Biological Bulletin*, 198: 225–244.

- Yan T., Zhou M.J., Fu M., Wang Y.F., Yu R.C., Li J.(2001). Inhibition of egg hatching success and larvae survival of the scallop, *Chlamys farreri*, associated with exposure to cells and cell fragments of the dinoflagellate *Alexandrium tamarense*. *Toxicon*, 39: 1239–1244.

- Yang W., Devaiah S.P., Pan X., Isaac G., Welti R., Wang X. (2007). AtPLAI is an acyl hydrolase involved in basal jasmonic acid production and *Arabidopsis* resistance to *Botrytis cinerea*. *J. Biol. Chem.*, 282: 282 (25).

- Young M.E., Keegstra K. and Froehlich J.E. (1999). GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. *Plant Physiol.*, 121: 237–243.